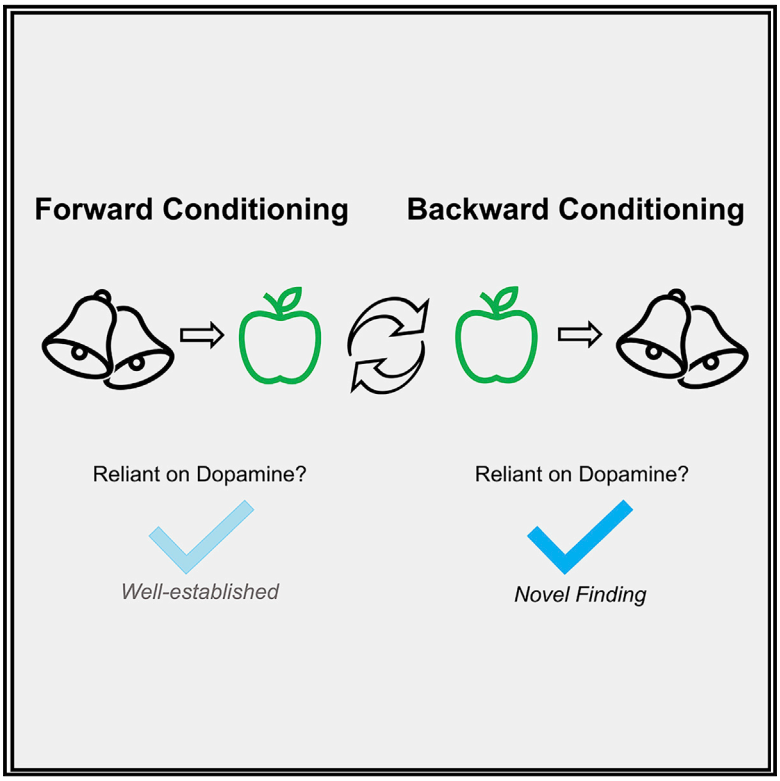
Report

Current Biology

Dopamine errors drive excitatory and inhibitory components of backward conditioning in an outcome-specific manner

Graphical abstract



Authors

Benjamin M. Seitz, Ivy B. Hoang, Lauren E. DiFazio, Aaron P. Blaisdell, Melissa J. Sharpe

Correspondence

melissa.j.sharpe@psych.ucla.edu

In brief

Seitz et al. show that VTA_{DA} neurons are necessary for the inhibitory and excitatory associations that develop during backward reward-cue conditioning. This poses a serious challenge to the value hypothesis of dopamine. Instead, these data support an emerging view that dopamine errors act as a teaching signal to drive associations between events.

Highlights

- Backward reward-cue conditioning produces inhibitory and excitatory associations
- Inhibition of VTA_{DA} neurons at reward-cue transition abolishes these associations
- Control studies show the backward cue did not become aversive or non-salient
- Suggests dopamine errors act as a general teaching signal, not a value signal

Seitz et al., 2022, Current Biology *32*, 3210–3218
July 25, 2022
https://doi.org/10.1016/j.cub.2022.06.035

CellPress

CellPress

Current Biology

Report

Dopamine errors drive excitatory and inhibitory components of backward conditioning in an outcome-specific manner

Benjamin M. Seitz,[1] Ivy B. Hoang,[1] Lauren E. DiFazio,[1] Aaron P. Blaisdell,[1] and Melissa J. Sharpe[1,2,3,*]
[1]Department of Psychology, University of California, Los Angeles, Portola Plaza, Los Angeles, CA 91602, USA
[2]Twitter: @MelissaJSharpe
[3]Lead contact
*Correspondence: melissa.j.sharpe@psych.ucla.edu
https://doi.org/10.1016/j.cub.2022.06.035

SUMMARY

For over two decades, phasic activity in midbrain dopamine neurons was considered synonymous with the prediction error in temporal-difference reinforcement learning.[1–4] Central to this proposal is the notion that reward-predictive stimuli become endowed with the scalar value of predicted rewards. When these cues are subsequently encountered, their predictive value is compared to the value of the actual reward received, allowing for the calculation of prediction errors.[5,6] Phasic firing of dopamine neurons was proposed to reflect this computation,[1,2] facilitating the backpropagation of value from the predicted reward to the reward-predictive stimulus, thus reducing future prediction errors. There are two critical assumptions of this proposal: (1) that dopamine errors can only facilitate learning about scalar value and not more complex features of predicted rewards, and (2) that the dopamine signal can only be involved in anticipatory cue-reward learning in which cues or actions precede rewards. Recent work[7–15] has challenged the first assumption, demonstrating that phasic dopamine signals across species are involved in learning about more complex features of the predicted outcomes, in a manner that transcends this value computation. Here, we tested the validity of the second assumption. Specifically, we examined whether phasic midbrain dopamine activity would be necessary for backward conditioning—when a neutral cue reliably *follows* a rewarding outcome.[16–20] Using a specific Pavlovian-to-instrumental transfer (PIT) procedure,[21–23] we show rats learn both excitatory and inhibitory components of a backward association, and that this association entails knowledge of the specific identity of the reward and cue. We demonstrate that brief optogenetic inhibition of VTA_{DA} neurons timed to the transition between the reward and cue reduces both of these components of backward conditioning. These findings suggest VTA_{DA} neurons are capable of facilitating associations between contiguously occurring events, regardless of the content of those events. We conclude that these data may be in line with suggestions that the VTA_{DA} error acts as a universal teaching signal. This may provide insight into why dopamine function has been implicated in myriad psychological disorders that are characterized by very distinct reinforcement-learning deficits.

RESULTS AND DISCUSSION

Early on, studies of associative learning were primarily concerned with understanding the basic mechanisms by which two events—broadly defined—become linked in the brain.[24,25] It is only recently that a shift has occurred such that major emphasis has been placed on the very specific temporal scenario in which a cue precedes a motivationally significant outcome (e.g., reward or pain).[5,6,26–28] Focusing on anticipatory cue→reward learning is advantageous in terms of computational modeling,[5,29–31] but it leaves many learning phenomena that do not involve this specific temporal order unexplained.[32]

An example of this trend relates to discovery of the dopamine prediction error. Shortly after it was revealed that dopamine neurons in the midbrain exhibit phasic signals to unexpected rewards,[1] this error signal was interpreted as being governed by computational rules that calculate scalar values in the context of anticipatory cue-reward learning.[1–4] Consequently, the study of the dopamine prediction error was almost exclusively focused on procedures involving anticipatory cue-reward associations that manipulate scalar value.[33–41] Only recently have we begun to explore the role of dopamine neurons in more complex paradigms outside of simple cue→reward learning. This work has uncovered that the prediction-error signal is capable of driving anticipatory learning of sensory events that transcend scalar value inherent in rewards, such as an association between two neutral cues.[7–15] Such findings question the assumption that dopamine neurons are "specialized" for anticipatory reward learning specifically, and whether anticipatory reward learning is "special" more generally.

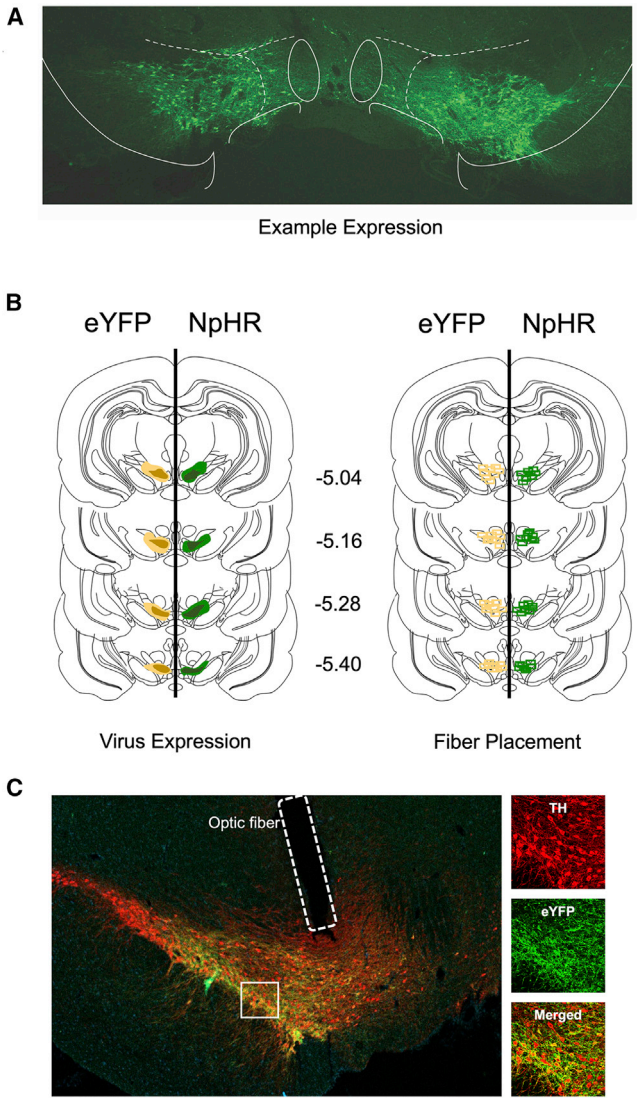


Figure 1. Histological representation of virus expression and fiber placement in TH-Cre rats

(A) Neurons in VTA expressing eYFP.

(B) Unilateral representation of the bilateral virus expression (left) and fiber placements (right). Fiber implants (green and yellow squares) were localized in the vicinity of NpHR (green) and eYFP (yellow) expression in VTA.

(C) High colocalization of TH and NpHR expression in VTA cell bodies (~91%, with example fiber placement).

Backward conditioning—when a reward is *followed* by a cue (reward→cue)—breaks this temporal mold and provides a serious challenge to current computational hypotheses of dopamine function. Backward conditioning cannot be explained by anticipatory cue-reward learning[42] as this learning can occur after a single reward-cue pairing[43] and can result in both excitatory and inhibitory associations.[16–20] That is, a backward cue is capable of exciting or inhibiting a representation of associated rewards, which motivates the animal toward or away from that specific reward. Here, we tested the necessity of dopamine transients in backward conditioning using an established procedure that combines backward conditioning with Pavlovian-to-instrumental transfer (PIT),[21–23] which probes for both the specific excitatory and inhibitory components of the association (Figure S1). This allows us to test whether dopamine neurons are exclusively involved in anticipatory cue→reward learning, or whether they function as a teaching signal to drive the formation of associations in a broader sense, regardless of whether those associations are anticipatory or backward, inhibitory or excitatory, and in a manner that transcends scalar value.

Inhibition of VTA_{DA} transients during backward conditioning prevents backward cues from exerting control over instrumental behavior

Rats expressing Cre-recombinase under the control of the tyrosine hydroxylase (TH) promoter (Rat Resource and Research Center [RRRC], Missouri)[44] received bilateral injections of either the inhibitory halorhodopsin (NpHR, AAV5-Ef1a-DIO eNpHR3.0-eYFP, n = 9) or a control virus that lacks the inhibitory opsin (eYFP, AAV5-Ef1a-DIO-eYFP, n = 9) in VTA (Figures 1A and 1B). Immunohistochemical verification showed a high degree of colocalization between Cre-dependent NpHR expression and TH in VTA (123 [eYFP+TH+]/135 [eYFP+]; n = 2; ~91%; Figure 1C). Optic fibers were also implanted bilaterally over VTA. After recovery, rats were food restricted and then received backward training, where two distinct rewards (pellets and maltodextrin solution) were each followed by one of two auditory cues (white noise and clicker [counterbalanced]; 8 days, 24 presentations per day). The pairing of the reward and cue was arranged such that the cue would be presented 10 s after the rat entered the magazine to consume the reward. This ensured the cue would be delivered shortly after the rats had consumed the reward. We delivered green light (532 nm, 16–18 mW output) into the VTA 500 ms before the onset of the cue and continuing for 2 s, as we have done previously.[13,45] We used these parameters to prevent phasic firing at the onset of the backward cue, which suppresses a potential prediction error to the backward cue, without producing a negative prediction error.[9,46] While a prediction error to the backward cue is not predicted by the scalar value account of dopamine, if these neurons are in fact facilitating the learning of two contiguously occurring events, inhibiting dopamine at cue onset should disrupt this learning.

Responding to the cues decreased over the course of conditioning, in line with other backward conditioning reports,[21–23,47] and this was similar across groups (Figure 2A; day, $F_{7,112}$= 4.593, p = 0.005; group, $F_{1,16}$ = 0.218, p = 0.647; day × group, $F_{7,112}$= 0.445, p = 0.741; Figure 2A). Rats then learned to press different levers for the distinct rewards (e.g., left lever→pellets, right lever→maltodextrin solution; counterbalanced), on an increasingly lean random-ratio schedule (CRF, RR5, RR10). All rats acquired the lever-pressing responses with no between-group differences (Figure 2B; day, $F_{7,112}$ = 650.415, p < 0.001; group, $F_{1,16}$ = 0.016, p = 0.901; day × group, $F_{7,112}$ = 1.521, p = 0.227; Figure 2B).

Finally, rats received a probe test in which both levers were available with no rewards delivered, and the backward cues were presented individually (i.e., the PIT test). The PIT test allows us to examine the nature of the associations that have developed during Pavlovian training. In standard PIT procedures with forward conditioning and two outcomes, an excitatory cue can elevate instrumental responding toward receiving the same

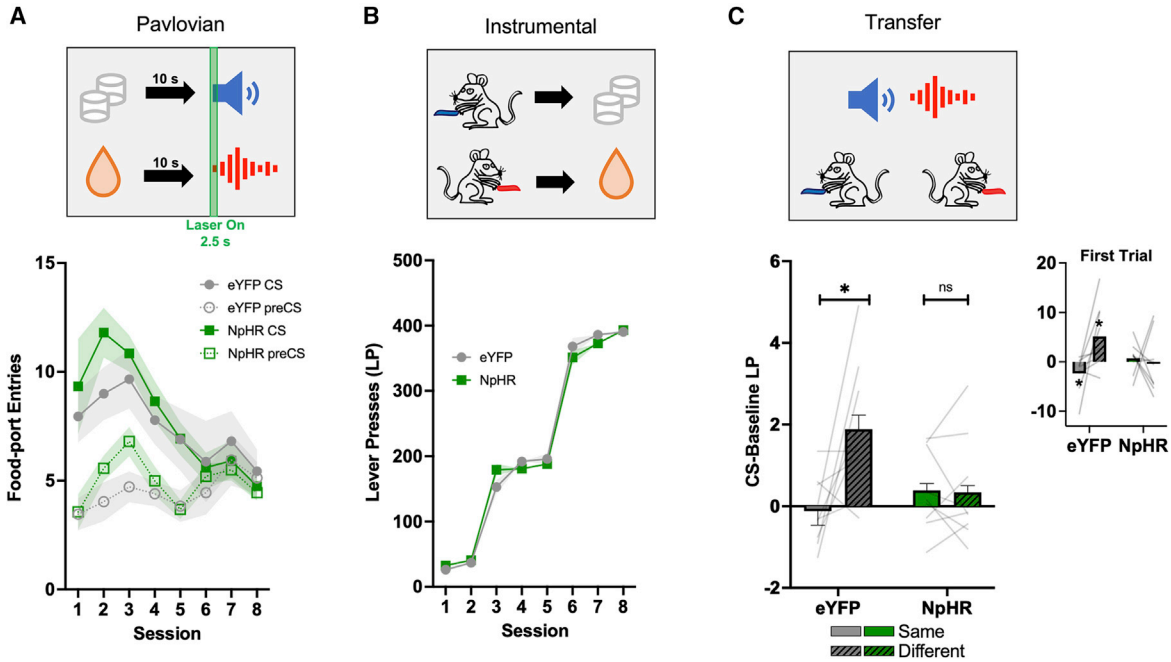


Figure 2. Inhibition of VTA_{DA} transients during backward conditioning prevents backward cues from exerting excitatory and inhibitory control over instrumental behavior

Rates of responding are represented as the number of entries into the food port or lever presses during cue presentation (±SEM), with lines indicating individual data points.

(A) Rats first learned backward relationships between two distinct rewards and two auditory cues (conditioned stimuli, CSs). The backward cue was presented 10 s after the rats entered the magazine to consume the rewards. Here, green light was delivered into VTA at the onset of the backward cue for 2.5 s to suppress phasic firing of dopamine neurons without producing a negative prediction error.[27] Responding during the cues decreased over the course of conditioning with no difference between groups.

(B) Rats then learned to make a left lever press to obtain one reward, and a right lever press to obtain the other. All rats acquired the instrumental responses for the rewards, with no difference between groups.

(C) Finally, during the PIT test, both levers were made available and the cues were individually presented without rewards (right). During the PIT test, the backward cues biased our eYFP group's responding away from the associated reward and toward the lever associated with the different reward (baseline, 1.2 [±0.8]; same, 1.1 [±0.7]; different, 3.1 [±2.1]). However, our NpHR group showed no change in responding from baseline during cue presentation or bias between the levers (baseline, 1.6 [±1.0]; same, 2.0 [±1.3]; different, 1.9 [±1.6]).

*$p < 0.05$, mean (±SEM). See also Figure S2.

outcome and not the different outcome (i.e., specific PIT), or elevate responding for both (i.e., general PIT).[40–42] However, inhibitory cues have repeatedly been shown to elevate responding specifically for the different outcome in PIT.[40–42] The exact reasoning for this is unclear. One possibility is that during instrumental conditioning, animals learn that a specific lever earns them one outcome and explicitly not the other outcome. Thus, when an inhibitory cue (like that produced by backward conditioning) is played at test, it indicates to the animal they will not receive one outcome and so they are invigorated to press for the other outcome (i.e. a specific excitation for the alternate outcome).[40–42] Another explanation is simply that presentation of an inhibitory cue generally excites behavior toward any alternative action, resulting in pressing a lever for the other outcome (i.e., general excitation for any other outcome). This remains to be tested, but it is clear that inhibitory associations have an outcome-specific influence on PIT that can be revealed by higher responding on the different lever and lower responding on the same lever.[40–42,45]

We calculated the difference in lever pressing during the cue, relative to baseline responding (average of same/different presses immediately before the cue).[40–42] Our results were in line with previous findings showing that the backward cue inhibits responding on the same lever and increases responding on the different lever. Specifically, in our eYFP group, backward cues biased lever-pressing away from the associated reward and toward the different reward (Figure 2C; lever × group, $F_{1,16} = 7.054$, $p = 0.017$; simple effect of lever, $F_{1,16} = 8.318$, $p = 0.020$; see Figure S2 for baseline responding). That is, the pellet-associated backward cue led to rats pressing more for solution, and the solution-associated backward cue led rats to press more for the pellet. This shows that the backward cues excite one behavior (lever press for different reward), while also inhibiting the other (lever press for same reward), in a sensory-specific manner. This difference in responding on the same versus different lever may indicate that the backward association is inhibitory and yet capable of exciting alternative responses[40–42] or that the backward association contains both excitatory and inhibitory properties for specific rewards.[20] Indeed, on the first trial, responding in our eYFP group to the different lever was significantly elevated from baseline ($t_8 = 2.474$, $p = 0.038$), whereas analyses suggested responding on the same lever was lower than baseline ($t_8 = 5.500$,

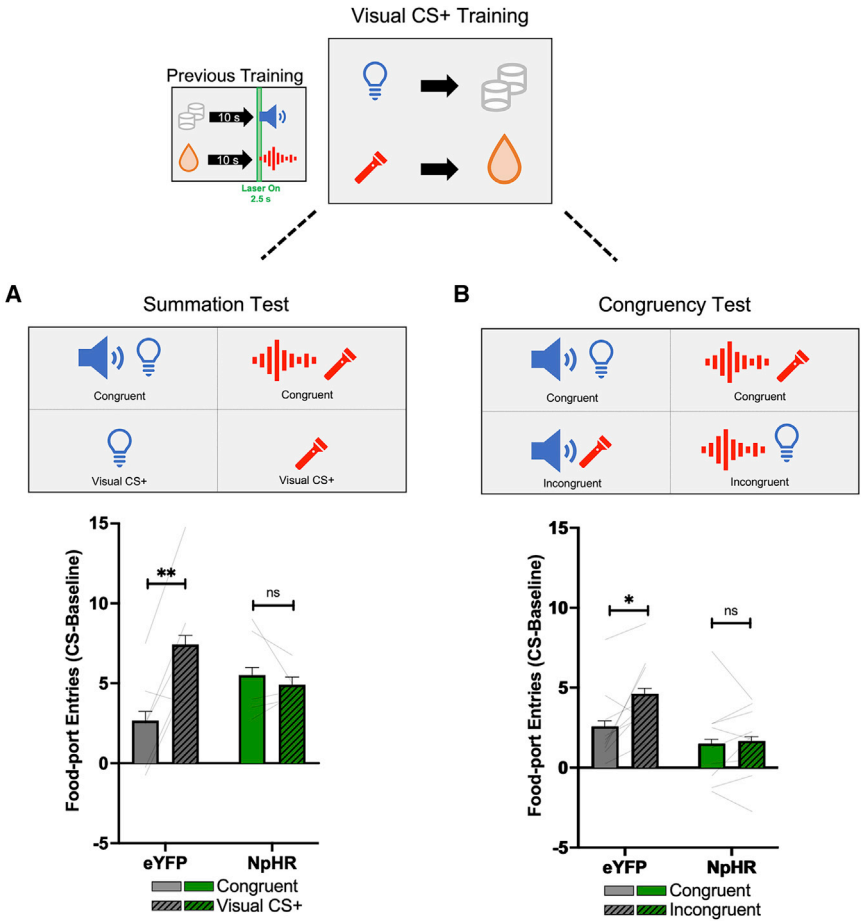


Figure 3. Inhibition of VTA_{DA} transients prevents backward cues from generally and specifically inhibiting Pavlovian responses

Responding is represented as number of entries into the food port during cue presentation (±SEM), with lines indicating individual data points. Top: visual forward training: to assess the nature of the deficit in the instrumental PIT test, we trained rats with two new forward cue-reward associations with visual stimuli (Figure S3). This allowed us to perform a number of tests with novel audiovisual compounds to investigate the source of the deficit in our NpHR group.

(A) Summation test: we tested responding to the visual cue by itself, relative to when it was presented in compound with the backward cue associated with the same outcome (i.e., congruent compound). If the backward cue is inhibitory, responding should be reduced on congruent trials relative to trials with the visual cue alone. Indeed, this is what we observed in the eYFP group. In contrast, the NpHR group showed the same high levels of responding to the visual cue whether or not it was presented in compound with the backward cue.

(B) Congruency test: the previous test indicates the backward cues are inhibitory when paired with the same outcome but did not test whether those cues possess specific or general inhibitory properties. To test this, we presented the visual cues in compound with the auditory cue predicting the same (congruent) or different (incongruent) reward. In the eYFP group, rats responded less on congruent relative to incongruent trials, suggesting the backward cues were specifically inhibitory. Again, there was no effect of the backward cues on responding to the visual cues in the NpHR group.

*$p < 0.05$, **$p < 0.01$. See also Figure S3.

$p = 0.050$). However, rats in our NpHR group showed no bias on lever responding and were not *elevated or decreased* from baseline lever-press responses (simple effect of lever, $F_{1,16} = 0.021$, $p = 0.889$; different lever versus baseline on first trial, $t_8 = 0.202$, $p = 0.845$; same lever versus baseline on first trial, $t_8 = 0.669$, $p = 0.504$). Finally, baseline lever press responding did not statistically differ between the two groups ($t_{16} = 0.946$, $p = 0.358$) (Figure S2A), and head entries into the food port did not differ between groups ($t_{16} = 0.480$, $p = 0.638$; Figure S2B). These findings suggest that inhibition of VTA_{DA} neurons at cue onset prevents the backward cues from exerting any effect over instrumental responding for the paired rewards, in an inhibitory or excitatory manner.

Inhibition of VTA_{DA} neurons prevents acquisition of the specific and general inhibitory components of backward conditioning

There are multiple interpretations that could be made from the failure of our NpHR group to use the backward cues to modulate instrumental performance. We suggest that VTA_{DA} inhibition prevented learning about the excitatory and inhibitory relationships between the rewards and backward cues. However, it is also possible that the NpHR rats still learned the inhibitory associations, but that the cues lacked some aspect of motivational significance that would allow them to exert control over an instrumental response. Further, a second interpretation of the PIT data is that the NpHR rats may have learned the backward cues were generally inhibitory of rewards. Thus, the performance of the NpHR rats during the PIT test could be interpreted as blanket inhibition of both lever-press responses during the PIT test—though this is unlikely as these rats did not reduce lever-pressing from baseline in the PIT test (Figure 2C).

To dissociate these accounts, we next taught the same rats two new forward associations with visual cues (e.g., house light→pellets; flashing light→maltodextrin solution; Figure S2). Training these new associations allowed us to investigate the impact of the backward cues on Pavlovian responding when presented in compound with the visual cues in an un-rewarded test session (i.e., a summation test). That is, when presented by themselves the visual cues should elicit high levels of responding because they signal the occurrence of a rewarding outcome. However, when each visual cue is presented in compound with the backward cue that signals the absence of the same outcome (i.e., a congruent compound), responding should be considerably reduced if the auditory cues are inhibitory. As predicted, responding in group eYFP was high when the visual cue was presented individually, while pairing it with the congruent backward cue significantly attenuated responding (Figure 3; summation test; cue type × group, $F_{1,9} = 11.893$, $p = 0.007$; simple effect of cue type, $F_{1,9} = 16.975$, $p = 0.009$). However, in the NpHR

group, the presence of the backward cue had no impact on responding to the visual cue (simple effect of cue type, $F_{1,9}$ = 0.375, p = 0.573). Not only do these cues pass the summation test, but other reports using an identical backward procedure have found that these cues also pass the retardation test.[22,47] Passing both tests is the gold standard for confirming inhibition[48] and suggests that inhibition of VTA_{DA} neurons prevented backward cues from acquiring inhibitory properties.

While the summation test above shows that VTA_{DA} inhibition prevents animals from learning the inhibitory component of backward cues in a Pavlovian procedure, they cannot speak to whether the backward cues generally or specifically inhibit Pavlovian responding in either the NpHR or eYFP rats. This is because we only presented a compound where both cues were associated with the same outcome and thus do not know if a backward cue presented in compound with a visual cue associated with the different outcome would similarly inhibit responding in a general fashion. A congruency test was used to tease apart the general versus specific nature of the inhibitory relationship that our NpHR group failed to learn. Specifically, just as we had presented in compound backward and forward cues associated with the same outcome (i.e., congruent compound), we could also present them in compound with backward and forward cues associated with different outcomes (i.e., incongruent). If the inhibitory relationship is specific, congruent compounds should show reduced responding relative to incongruent compounds. However, if the inhibitory relationship is general, there should be no difference between congruent and incongruent compounds. In our eYFP group, we observed a reduction in responding on congruent relative to incongruent compound trials (Figure 3; congruency test; compound × group, $F_{1,16}$ = 4.571, p = 0.048; simple effect of compound, $F_{1,16}$ = 8.790, p = 0.018). In contrast, rats in group NpHR showed no difference in Pavlovian responding during congruent versus incongruent trials (simple effect of compound, $F_{1,16}$ = 0.096, p = 0.765), confirming they had not learned the specific inhibitory associations with the backwards cue, and it was not a more general deficit in using the Pavlovian cues to exert control over instrumental behavior. Note that while responding in group NpHR appears low in the congruency test, potentially suggesting that general inhibition remains intact in NpHR rats, if we compare NpHR responding to the forward visual cues alone on the day of training prior to the congruency test Figure S2 with responding to the congruent and incongruent compound at test, there is no difference (Figure S2; $F_{2,16}$ = 0.779, p = 0.476). Thus, it appears most likely that rats in our NpHR group responded to the visual stimuli as they had in training, and when those cues were paired with the backward cues responding was unaffected. This is in clear contrast to the impact of responding that the backward cues had when paired with forward visual cues in our eYFP group.

Inhibition of VTA_{DA} neurons at cue onset in forward conditioning does not prevent learning or make cues aversive

Our present results showed that brief optogenetic inhibition of VTA_{DA} neurons at cue onset in backward conditioning prevented rats from learning the excitatory and inhibitory components in backward conditioning, which we would interpret as indicating the dopamine prediction error is a broad teaching signal that transcends both scalar value and anticipatory cue-reward associative structures. However, it is possible that inhibiting VTA_{DA} neurons at cue onset somehow made these cues aversive, or simply reduced their salience so that they could not be learned about. Indeed, dopamine neurons often fire to novel cues,[47] and a recent report showed preventing dopamine release at cue onset can decelerate learning, whereas stimulating cue-evoked phasic dopamine firing can accelerate learning.[48] This could suggest that phasic dopamine activity at cue onset signals the cue with an inherent value or salience. According to this view, we may have altered the significance of the cue through our inhibition of VTA_{DA} at cue onset rather than disrupting the association of the backward cue with the reward, per se. In order to rule out this interpretation, we inhibited VTA_{DA} neurons at cue onset during forward conditioning. If inhibition of VTA_{DA} neurons at cue onset during backward conditioning made these cues aversive or reduced their salience to the point that they could not be learned about, inhibition of VTA_{DA} neurons at cue onset in forward conditioning should similarly wipe out learning. It is worth noting that a role for VTA_{DA} inhibition at the onset of the antecedent (in forward conditioning, at cue onset) in preventing learning would also be against recent findings demonstrating that inhibition of VTA_{DA} at cue onset during the blocking procedure does not prevent blocking,[45] suggesting that the dopaminergic signal produced by the cue does not contain reward prediction (or "cue significance") but rather an error in the expected versus experienced events (see Maes et al.[45] and Schultz[49] for more discussion). Taken together, this is consistent with a view that the predominant role of dopamine neurons is acting as a teaching signal and not one that contains an upcoming prediction or saliency of the cue.

To test whether inhibition of VTA_{DA} neurons at cue onset would prevent forward conditioning, we taught all rats new forward relationships between two novel auditory cues (siren and tone) and two distinct food rewards. Rats learned these new relationships in a novel context so as to prevent potential carry-over effects from the previous studies. We delivered green light (532 nm, 16–18 mW output) to VTA_{DA} neurons at cue onset for one of the auditory cues but not the other (counterbalanced), using the same inhibition parameters as backward conditioning (i.e., 2.5 s inhibition at cue onset). We observed no difference in acquisition between the cue with laser on versus the cue with the laser off in either group (Figure 4A; day, $F_{7,112}$ = 2.741, p = 0.060; laser, $F_{1,16}$ = 0.947, p = 0.345; group, $F_{1,16}$ = 0.079, p = 0.782; day × group, $F_{7,112}$ = 0.246, p = 0.845; day × laser, $F_{7,112}$ = 1.266, p = 0.291; laser × group, $F_{1,16}$ = 2.051, p = 0.171; day × laser × group, $F_{7,112}$ = 0.522, p = 0.734). However, in the NpHR group, the cue with the laser on showed a small, but statistically non-significant, retardation of acquisition (simple effect of laser status, $F_{1,16}$ = 3.940, p = 0.082; Figure 4A), approximately replicating the results of Morrens et al.[50] Despite this, responding during the two cues was virtually indistinguishable after the initial sessions, and an extinction test after the completion of training revealed no between-group or within-group differences in responding (Figure 4B; laser status, $F_{1,16}$ = 0.236, p = 0.634; group, $F_{1,16}$ = 0.011, p = 0.916; laser status × group, $F_{1,16}$ = 0.006, p = 0.937). These results suggest that VTA_{DA} inhibition at cue onset does not prevent learning about the cue-reward association—although it modestly

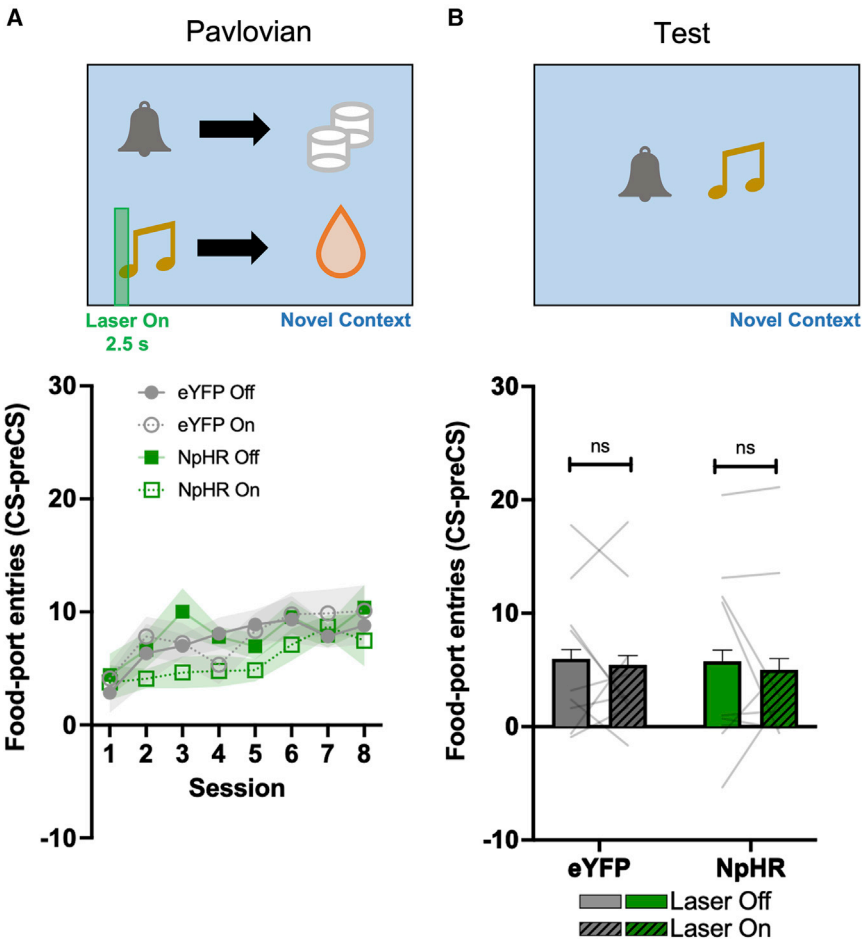


Figure 4. Inhibition of VTA_{DA} transients at cue onset in forward conditioning does not impair learning

Responding is represented as number of entries into the food port during cue presentation (±SEM). To ensure our findings could not be the result of VTA_{DA} inhibition at cue onset causing the backward cues to become aversive or significantly reducing their salience, we taught rats novel auditory cue-reward associations with VTA_{DA} inhibition at cue onset.

(A) Rats learned forward relationships with two novel auditory cues, one of which received light delivery into VTA at cue onset. Pavlovian training progressed normally for both cues in group eYFP, with a non-significant reduction in responding to the NpHR group at the beginning of training.

(B) We then tested responding to the auditory cues by themselves without laser inhibition. There were no differences in responding between groups, or between cues. These results suggest VTA_{DA} inhibition at cue onset does not prevent learning.

decelerates learning. Thus, the results from our backward training cannot be explained by VTA_{DA} neuronal inhibition reducing the salience of the cues to the extent that they cannot be learned about, or by making them in some way aversive.

Conclusion

These data show that backward reward→cue associations can modulate instrumental behavior in an excitatory, inhibitory, and outcome-specific manner. Further, inhibition of VTA_{DA} neurons at the onset of the backward cue to suppress phasic firing of dopamine neurons prevents learning of the inhibitory and excitatory sensory-specific backward associations. We also ruled out the possibility that inhibiting VTA_{DA} neurons at cue onset simply prevents learning by reducing cue salience or making cues aversive. These data are consistent with recent work implicating phasic activity in VTA_{DA} neurons in learning outside the context of scalar values,[7–15] and extend this research in critical ways.

Canonical models[1–4,31,49] of the dopamine prediction error have restricted these neurons to anticipatory cue-reward learning, via the backpropagation of scalar value to reward-predictive cues. However, our data show that VTA_{DA} transients are necessary for the excitatory and inhibitory components of backward conditioning in a manner that entails specific knowledge of the identity of the events. This comes at a time when there is mounting evidence that the dopamine error facilitates far more complex learning than that afforded by the backpropagation of scalar value.[51,52] For example, VTA_{DA} transients are necessary and sufficient for learning associations between two neutral cues (e.g., tone→light), and VTA_{DA} neurons achieve this without making the neutral cues valuable in and of themselves.[9,10,13,53] Similarly, artificially inducing dopamine prediction errors during cue-reward learning subsequently allows the cue to evoke a detailed representation of the reward.[51] Results like these and others,[25–30] consistent with those reported here, suggest VTA_{DA} neurons are capable of producing an error that facilitates "model-based" learning, which refers to an ability to associate (and predict) sensory representations of events. However, even an error signal that facilitates model-based learning cannot fully explain our results with backward conditioning. This is because model-based accounts still ultimately rely on value backpropagating to earlier predictors of reward, albeit in the context of more complex associative structures, whether inferred or directly experienced.[29,54]

One potential limitation of our study is the continued use of the same animals over repeated experiments. This may have resulted in carry-over effects that influenced responding in the later experiments. However, the results from our final forward control experiment are essentially a replication of Morrens et al.[46] Further, using the same rats for both these experiments also ensured the manipulation of the same dopamine neurons in the same animals could produce these dissociable effects in backward and forward training. Another limitation is that we have found VTA_{DA} neurons to be necessary for backward conditioning but have not characterized their normal firing pattern during this procedure. Recording these neurons during backward conditioning would be especially interesting and should be pursued in future research. Backward conditioning has been considerably understudied in learning theory and behavioral

neuroscience, and further investigation into its neural substrates represents an exciting opportunity given its theoretical importance and our results indicating dopamine involvement.

How should we interpret the necessity of VTA_{DA} neurons in backward conditioning? The most parsimonious explanation of our data and other recent findings is that VTA_{DA} neurons are computing prediction errors between contiguously occurring events. Thus, regardless of whether the events are two contiguously occurring cues (as in sensory preconditioning[13] and second-order conditioning[45]) or other sensory events, VTA_{DA} neurons might be sending errors that reflect a mismatch between sensory expectations and events. That is, it could be considered a more general sensory prediction error that serves to reduce the presence of prediction errors in our everyday sensory experience, which sometimes involves events that possess value (like rewards). It should be noted in our procedure there is a 10 s gap between the two events. Thus, how far two events can be spaced apart to still be considered contiguous is unclear. However, neither lengthy temporal gaps nor the backward arrangement precludes learning and is theoretically accounted for in reinforcement-learning models.[55] Indeed, the original Rescorla-Wagner model,[6] which serves as the basis for temporal difference reinforcement learning (TDRL) algorithms, is agnostic toward whether prediction errors are value-based or more cognitive like we are now suggesting. Such a stance would argue that VTA_{DA} neurons are contributing to learning in ways more closely aligned with historical interpretations of associative learning[25] and less with modern TDRL-centric interpretations. While viewing dopamine neurons in this light enables them to be involved in far more complex forms of learning, there are some learning phenomena that dopamine's involvement is not particularly clear. This is due in part to contradictory data (e.g., heterogeneity of dopamine neuronal responding during aversive procedures). This may be due to different dopamine subpopulations having different functions, coupled with their distinct projection profiles. Thus, trying to establish "the" role of dopamine in learning and behavior might be futile. Indeed, dopamine neurons do not act as a monolith,[16,56,57] and our approach of broadly inhibiting these neurons does not necessarily imply all VTA_{DA} neurons contribute to backward conditioning.[58]

While exact boundaries of the types of associations these neurons contribute to have yet to be established, the implications that a proportion of dopamine neurons could be acting as a more universal teaching signal are profound. First, if these neurons contribute to mentally linking contiguously occurring events, rather than predicting rewards (either proximally or distally), it would explain why they have been found to be necessary for higher-order conditioning,[31,44] and also place dopamine at the center of many complex forms of cognition (e.g., spatial and causal reasoning).[59] Ultimately, this may have important implications in pathologies characterized by abnormal dopaminergic functioning (e.g., schizophrenia and addiction). Indeed, an excess of subcortical dopamine (a trademark of schizophrenia) would be expected to be correlated with an excess in learning relationships between potentially irrelevant events—which could result in hallucinogenic or delusional experiences.[56,57,60–63] To expand, not all co-occurring events need be associated, and there are also regions (e.g., lateral hypothalamus) whose function appears to be opposing the learning of relationships that do not immediately predict rewards.[64,65] Such findings situate the VTA_{DA} prediction error at the center of a dynamic system whose main function is to direct learning in one way or another via distinct circuits, depending on current context or motivational state, and past experience. Future research will tell how far we can push the boundaries of dopamine's involvement in learning and cognition.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Subjects
- METHOD DETAILS
  - Surgeries
  - Apparatus
  - Backward Pavlovian training
  - Instrumental training
  - Transfer test
  - Forward conditioning with visual cues
  - Congruency test
  - Summation test
  - VTA_{DA} neuronal inhibition at cue onset in forward conditioning
  - Histology
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.06.035.

ACKNOWLEDGMENTS

This work was supported by an NSF CAREER (NSF2143910), a BBRF award (BBRF30637), and a Society of Hellman Fellows award to M.J.S. This work was also supported by NSF grant BCS-1844144 to A.P.B. and NSF grant DGE-1650604 to B.M.S.

AUTHOR CONTRIBUTIONS

B.M.S., A.P.B., and M.J.S. designed the experiments. B.M.S., I.B.H., and L.E.D. conducted the experiments. B.M.S. analyzed the data. B.M.S., A.P.B., and M.J.S. interpreted the data. B.M.S. and M.J.S. wrote the paper with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We worked to ensure sex balance in the selection of non-human subjects. One or more of the authors of this paper self-identifies as living with a disability. The author list of this paper includes contributors from the location where the research was conducted who participated in the data collection, design,

analysis, and/or interpretation of the work. While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list.

Received: January 26, 2022
Revised: April 29, 2022
Accepted: June 13, 2022
Published: June 24, 2022

REFERENCES

1. Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599. https://doi.org/10.1126/science.275.5306.1593.
2. Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. Nature *412*, 43–48. https://doi.org/10.1038/35083500.
3. Glimcher, P.W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. Proc. Natl. Acad. Sci. USA *108*, 15647–15654. https://doi.org/10.1073/pnas.1014269108.
4. Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. Nat. Rev. Neurosci. *17*, 183–195. https://doi.org/10.1038/nrn.2015.26.
5. Sutton, R.S., and Barto, A.G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. Psychol. Rev. *88*, 135–170. https://doi.org/10.1037/0033-295x.88.2.135.
6. Rescorla, R.A., and Wagner, A.R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In Classical Conditioning II: Current Research and Theory, A. Black, and W. Prokasy, eds. (Appleton-Centrury-Crofts), pp. 64–99.
7. Stalnaker, T.A., Howard, J.D., Takahashi, Y.K., Gershman, S.J., Kahnt, T., and Schoenbaum, G. (2019). Dopamine neuron ensembles signal the content of sensory prediction errors. eLife *8*, e49315. https://doi.org/10.7554/elife.49315.
8. Howard, J.D., and Kahnt, T. (2018). Identity prediction errors in the human midbrain update reward-identity expectations in the orbitofrontal cortex. Nat. Commun. *9*, 1611.
9. Sharpe, M.J., Batchelor, H.M., Mueller, L.E., Yun Chang, C., Maes, E.J.P., Niv, Y., and Schoenbaum, G. (2020). Dopamine transients do not act as model-free prediction errors during associative learning. Nat. Commun. *11*, 106. https://doi.org/10.1038/s41467-019-13953-1.
10. Sadacca, B.F., Jones, J.L., and Schoenbaum, G. (2016). Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. eLife *5*, e13665. https://doi.org/10.7554/elife.13665.
11. Takahashi, Y.K., Batchelor, H.M., Liu, B., Khanna, A., Morales, M., and Schoenbaum, G. (2017). Dopamine neurons respond to errors in the prediction of sensory features of expected rewards. Neuron *95*, 1395–1405.e3. https://doi.org/10.1016/j.neuron.2017.08.025.
12. Chang, C.Y., Gardner, M., Di Tillio, M.G., and Schoenbaum, G. (2017). Optogenetic blockade of dopamine transients prevents learning induced by changes in reward features. Curr. Biol. *27*, 3480–3486.e3. https://doi.org/10.1016/j.cub.2017.09.049.
13. Sharpe, M.J., Chang, C.Y., Liu, M.A., Batchelor, H.M., Mueller, L.E., Jones, J.L., Niv, Y., and Schoenbaum, G. (2017). Dopamine transients are sufficient and necessary for acquisition of model-based associations. Nat. Neurosci. *20*, 735–742. https://doi.org/10.1038/nn.4538.
14. Keiflin, R., Pribut, H.J., Shah, N.B., and Janak, P.H. (2019). Ventral tegmental dopamine neurons participate in reward identity predictions. Curr. Biol. *29*, 93–103.e3. https://doi.org/10.1016/j.cub.2018.11.050.
15. Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H.J., Ornelas, S., Koay, S.A., Thiberge, S.Y., Daw, N.D., Tank, D.W., and Witten, I.B. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. Nature *570*, 509–513. https://doi.org/10.1038/s41586-019-1261-9.
16. Barnet, R.C., and Miller, R.R. (1996). Second-order excitation mediated by a backward conditioned inhibitor. J. Exp. Psychol. Anim. Behav. Process. *22*, 279–296. https://doi.org/10.1037/0097-7403.22.3.279.
17. Prével, A., Rivière, V., Darcheville, J.C., Urcelay, G.P., and Miller, R.R. (2019). Excitatory second-order conditioning using a backward first-order conditioned stimulus: a challenge for prediction error reduction. Q. J. Exp. Psychol. *72*, 1453–1465. https://doi.org/10.1177/1747021818793376.
18. Chang, R.C., Blaisdell, A.P., and Miller, R.R. (2003). Backward conditioning: mediation by the context. J. Exp. Psychol. Anim. Behav. Process. *29*, 171–183. https://doi.org/10.1037/0097-7403.29.3.171.
19. Urushihara, K. (2004). Excitatory backward conditioning in an appetitive conditioned reinforcement preparation with rats. Behav. Processes *67*, 477–489. https://doi.org/10.1016/j.beproc.2004.08.002.
20. Cole, R.P., and Miller, R.R. (1999). Conditioned excitation and conditioned inhibition acquired through backward conditioning. Learn. Motiv. *30*, 129–156. https://doi.org/10.1006/lmot.1998.1027.
21. Laurent, V., Wong, F.L., and Balleine, B.W. (2017). The lateral habenula and its input to the rostromedial tegmental nucleus mediates outcome-specific conditioned inhibition. J. Neurosci. *37*, 10932–10942. https://doi.org/10.1523/jneurosci.3415-16.2017.
22. Laurent, V., and Balleine, B.W. (2015). Factual and counterfactual action-outcome mappings control choice between goal-directed actions in rats. Curr. Biol. *25*, 1074–1079. https://doi.org/10.1016/j.cub.2015.02.044.
23. Laurent, V., Wong, F.L., and Balleine, B.W. (2015). δ-Opioid receptors in the accumbens shell mediate the influence of both excitatory and inhibitory predictions on choice. Br. J. Pharmacol. *172*, 562–570. https://doi.org/10.1111/bph.12731.
24. Pavlov, I.P. (1927). Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex (Oxford University Press).
25. Bolles, R. (1993). The Story of Psychology: A Thematic History (Brooks/Cole Publishing Co.).
26. Mackintosh, N.J. (1975). A theory of attention: variations in the associability of stimuli with reinforcement. Psychol. Rev. *82*, 276–298. https://doi.org/10.1037/h0076778.
27. Pearce, J.M., and Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol. Rev. *87*, 532–552. https://doi.org/10.1037/0033-295x.87.6.532.
28. Kamin, L.J. (1969). Predictability, surprise, attention, and conditioning. In Punishment Aversive Behavior, B.A. Campbell, and R.M. Church, eds. (Appleton-Century Crofts), pp. 279–296.
29. Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. *8*, 1704–1711. https://doi.org/10.1038/nn1560.
30. Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. *36*, 181–204. https://doi.org/10.1017/s0140525x12000477.
31. Schultz, W., and Dickinson, A. (2000). Neuronal coding of prediction errors. Annu. Rev. Neurosci. *23*, 473–500. https://doi.org/10.1146/annurev.neuro.23.1.473.
32. Miller, R.R., Barnet, R.C., and Grahame, N.J. (1995). Assessment of the Rescorla-Wagner model. Psychol. Bull. *117*, 363–386. https://doi.org/10.1037/0033-2909.117.3.363.
33. Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature *482*, 85–88. https://doi.org/10.1038/nature10754.
34. Saunders, B.T., Richard, J.M., Margolis, E.B., and Janak, P.H. (2018). Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. Nat. Neurosci. *21*, 1072–1083. https://doi.org/10.1038/s41593-018-0191-4.
35. Hollerman, J.R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. Nat. Neurosci. *1*, 304–309. https://doi.org/10.1038/1124.
36. Chang, C.Y., Esber, G.R., Marrero-Garcia, Y., Yau, H.J., Bonci, A., and Schoenbaum, G. (2015). Brief optogenetic inhibition of dopamine neurons

mimics endogenous negative reward prediction errors. Nat. Neurosci. *19*, 111–116. https://doi.org/10.1038/nn.4191.

37. Steinberg, E.E., Keiflin, R., Boivin, J.R., Witten, I.B., Deisseroth, K., and Janak, P.H. (2013). A causal link between prediction errors, dopamine neurons and learning. Nat. Neurosci. *16*, 966–973. https://doi.org/10.1038/nn.3413.

38. Tsai, H.C., Zhang, F., Adamantidis, A., Stuber, G.D., Bonci, A., De Lecea, L., and Deisseroth, K. (2009). Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. Science *324*, 1080–1084. https://doi.org/10.1126/science.1168878.

39. Lak, A., Stauffer, W.R., and Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. Proc. Natl. Acad. Sci. USA *111*, 2343–2348. https://doi.org/10.1073/pnas.1321596111.

40. Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. Science *307*, 1642–1645. https://doi.org/10.1126/science.1105370.

41. Fiorillo, C.D. (2013). Two dimensions of value: dopamine neurons represent reward but not aversiveness. Science *341*, 546–549. https://doi.org/10.1126/science.1238699.

42. Spetch, M.L., Wilkie, D.M., and Pinel, J.P. (1981). Backward conditioning: a reevaluation of the empirical evidence. Psychol. Bull. *89*, 163–175. https://doi.org/10.1037/0033-2909.89.1.163.

43. Mahoney, W.J., and Ayres, J.J.B. (1976). One-trial simultaneous and backward fear conditioning as reflected in conditioned suppression of licking in rats. Anim. Learn. Behav. *4*, 357–362. https://doi.org/10.3758/bf03214421.

44. Witten, I.B., Steinberg, E.E., Lee, S.Y., Davidson, T.J., Zalocusky, K.A., Brodsky, M., Yizhar, O., Cho, S.L., Gong, S., Ramakrishnan, C., et al. (2011). Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. Neuron *72*, 721–733. https://doi.org/10.1016/j.neuron.2011.10.028.

45. Maes, E.J.P., Sharpe, M.J., Usypchuk, A.A., Lozzi, M., Chang, C.Y., Gardner, M.P.H., Schoenbaum, G., and Iordanova, M.D. (2020). Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors. Nat. Neurosci. *23*, 176–178. https://doi.org/10.1038/s41593-019-0574-1.

46. Chang, C.Y., Gardner, M.P.H., Conroy, J.C., Whitaker, L.R., and Schoenbaum, G. (2018). Brief, but not prolonged, pauses in the firing of midbrain dopamine neurons are sufficient to produce a conditioned inhibitor. J. Neurosci. *38*, 8822–8830. https://doi.org/10.1523/jneurosci.0144-18.2018.

47. Delamater, A.R., LoLordo, V.M., and Sosa, W. (2003). Outcome-specific conditioned inhibition in Pavlovian backward conditioning. Learn. Behav. *31*, 393–402. https://doi.org/10.3758/bf03196000.

48. Rescorla, R.A. (1969). Pavlovian conditioned inhibition. Psychol. Bull. *72*, 77–94. https://doi.org/10.1037/h0027760.

49. Schultz, W. (1998). Predictive reward signal of dopamine neurons. J. Neurophysiol. *80*, 1–27. https://doi.org/10.1152/jn.1998.80.1.1.

50. Morrens, J., Aydin, Ç., Janse van Rensburg, A., Esquivelzeta Rabell, J., and Haesler, S. (2020). Cue-evoked dopamine promotes conditioned responding during learning. Neuron *106*, 142–153.e7. https://doi.org/10.1016/j.neuron.2020.01.012.

51. Langdon, A.J., Sharpe, M.J., Schoenbaum, G., and Niv, Y. (2018). Model-based predictions for dopamine. Curr. Opin. Neurobiol. *49*, 1–7. https://doi.org/10.1016/j.conb.2017.10.006.

52. Sharpe, M.J., and Schoenbaum, G. (2018). Evaluation of the hypothesis that phasic dopamine constitutes a cached-value signal. Neurobiol. Learn. Mem. *153*, 131–136. https://doi.org/10.1016/j.nlm.2017.12.002.

53. Sharpe, M.J., Batchelor, H.M., and Schoenbaum, G. (2017). Preconditioned cues have no value. eLife *6*, e28362. https://doi.org/10.7554/elife.28362.

54. Gardner, M.P.H., Schoenbaum, G., and Gershman, S.J. (2018). Rethinking dopamine as generalized prediction error. Proc. R. Soc. B *285*, 20181645. https://doi.org/10.1098/rspb.2018.1645.

55. Wagner, A.R., and Brandon, S. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In Contemporary Learning Theories: Pavlovian Conditioning and the Status of Traditional Learning Theory, S. Klein, and R. Mowrer, eds. (Erlbaum), pp. 149–189.

56. Corlett, P.R., Taylor, J.R., Wang, X.J., Fletcher, P.C., and Krystal, J.H. (2010). Toward a neurobiology of delusions. Prog. Neurobiol. *92*, 345–369. https://doi.org/10.1016/j.pneurobio.2010.06.007.

57. Corlett, P.R., Murray, G.K., Honey, G.D., Aitken, M.R.F., Shanks, D.R., Robbins, T.W., Bullmore, E.T., Dickinson, A., and Fletcher, P.C. (2007). Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. Brain *130*, 2387–2400. https://doi.org/10.1093/brain/awm173.

58. Lee, R.S., Engelhard, B., Witten, I.B., and Daw, N.D. (2022). A vector reward prediction error model explains dopaminergic heterogeneity. Preprint at bioRxiv. https://doi.org/10.1101/2022.02.28.482379.

59. Seitz, B.M., Blaisdell, A.P., and Sharpe, M.J. (2021). Higher-order conditioning and dopamine: charting a path forward. Front. Behav. Neurosci. *15*, 745388. https://doi.org/10.3389/fnbeh.2021.745388.

60. Millard, S.J., Bearden, C.E., Karlsgodt, K.H., and Sharpe, M.J. (2021). The prediction-error hypothesis of schizophrenia: new data point to circuit-specific changes in dopamine activity. Neuropsychopharmacol. *47*, 628–640. https://doi.org/10.1038/s41386-021-01188-y.

61. Jensen, J., Willeit, M., Zipursky, R.B., Savina, I., Smith, A.J., Menon, M., Crawley, A.P., and Kapur, S. (2007). The formation of abnormal associations in schizophrenia: neural and behavioral evidence. Neuropsychopharmacol. *33*, 473–479. https://doi.org/10.1038/sj.npp.1301437.

62. Morris, R.W., Vercammen, A., Lenroot, R., Moore, L., Langton, J.M., Short, B., Kulkarni, J., Curtis, J., Curtis, J., O'Donnell, M., et al. (2012). Disambiguating ventral striatum fMRI-related BOLD signal during reward prediction in schizophrenia. Mol. Psychiatry *17*, 280–289. https://doi.org/10.1038/mp.2011.75.

63. Morris, R., Griffiths, O., Le Pelley, M.E., and Weickert, T.W. (2013). Attention to irrelevant cues is related to positive symptoms in schizophrenia. Schizophr. Bull. *39*, 575–582. https://doi.org/10.1093/schbul/sbr192.

64. Hoang, I.B., and Sharpe, M.J. (2021). The basolateral amygdala and lateral hypothalamus bias learning towards motivationally significant events. Curr. Opin. Behav. Sci. *41*, 92–97. https://doi.org/10.1016/j.cobeha.2021.04.014.

65. Sharpe, M.J., Batchelor, H.M., Mueller, L.E., Gardner, M.P.H., and Schoenbaum, G. (2021). Past experience shapes the neural circuits recruited for future learning. Nat. Neurosci. *24*, 391–400. https://doi.org/10.1038/s41593-020-00791-4.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Red (594) secondary antibodies	Thermofischer	https://www.thermofisher.com/antibody/product/Goat-anti-Rabbit-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-11012; RRID: AB_2534079
Bacterial and virus strains		
pAAV-Ef1a-DIO-eYFP (AAV5)	Addgene	https://www.addgene.org/27056/
pAAV-Ef1a-DIO eNpHR 3.0-eYFP (AAV5)	Addgene	http://www.addgene.org/26966/#26966-AAV5
Chemicals, peptides, and recombinant proteins		
Goat Serum	Millipore Sigma	https://www.emdmillipore.com/US/en/product/Normal-Goat-Serum-Lyophilized-Solid,EMD_BIO-566380#anchor_orderingcomp
DAPI wash	Thermofischer	https://www.thermofisher.com/order/catalog/product/D1306?SID=srch-srp-D1306
Prolong Antifade no DAPI	Thermofischer	https://www.thermofisher.com/order/catalog/product/P36930?SID=srch-hj-P36930
Deposited data		
Raw data and statistical analyses	Open Source Framework	https://osf.io/a2dpf/
Experimental models: Organisms/strains		
TH::Cre on Long-Evans Background	Rat Resource and Research Center (RRRC), Missouri	Karl Deisseroth[44]

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Melissa Sharpe (melissa.j.sharpe@psych.ucla.edu).

Materials availability

This study did not generate new reagents.

Data and code availability

- All data have been deposited at Open Science Framework (OSF) and are publicly available as of the date of publication. DOIs are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Subjects

18 transgenic Long-Evans rats (8 Female, 10 Male) expressing Cre-recombinase under the control of the tyrosine hydroxylase (TH) promoter (Rat Resource and Research Center) were used in this study.[43] Rats were randomly allocated to groups and matched for age and sex. Rats were maintained on a 12-h 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 the UCLA Institutional Animal Care and Use Committee.

METHOD DETAILS

Surgeries

Surgical procedures have been described elsewhere.[13] Briefly, rats received bilateral infusions of 1.0-2.0 μL of AAV5-EF1α-DIO-eYFP (n = 9) or eNpHR3.0-eYFP (n = 9) 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 Addgene. During surgery, optic fibers were implanted bilaterally (200-μm diameter, Thorlabs, 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

Behavioral sessions were conducted in identical sound-attenuated conditioning chambers (Med Associates, St. Albans, VT). The chambers contained 2 retractable levers that could be inserted to the left and right of a recessed food delivery port in the front wall when triggered. A photobeam entry detector was positioned at the entry to the food port. The chambers were also equipped with syringe pumps to deliver 15% maltodextrin solution in 0.1 ml increments through a stainless steel tube into a custom-designed well in the food port and a pellet dispenser to deliver a single 45-mg sucrose pellet (Bio-Serv, Frenchtown, NJ). Both a tone and white noise generator were attached to individual speakers on the wall opposite the lever and magazine. A 3-watt, 24-volt house light mounted on the top of the back wall opposite the food cup and two white lights were mounted above the levers and served as visual cues.

Backward Pavlovian training

Rats received 8 consecutive days of Pavlovian conditioning. Rewards (sucrose pellet or maltodextrin solution) were delivered into the food port, and auditory cues (clicker or white noise) were played 10 s following the first entry into the magazine. Reward-cue relationships were fully counterbalanced. Cue duration varied from 2-58 s with an average of 30 s. Data are presented as average entries per minute. Variable cue duration was chosen to stay consistent with the procedure described elsewhere[21–23] (Figure S1) and because variable cue length helps promote instrumental responding at test by preventing the animal from timing the delivery of the outcome. Stimuli were presented 12 times each in a pseudorandom order with a variable inter-trial-interval (ITI) ranging from 80-190 s with an average of 125 s. Approximately 30% of the reward-cue trials (e.g. pellet→white noise) were followed by the same trial type (i.e., pellet→white noise), which reduces the likelihood rats will associate the backward cue and the alternate reward (i.e., white noise→malto) in a forward excitatory manner. The variable ITI and pseudo-random presentation of the backward reward-cue trials also helps reduce the possibility of subjects learning an excitatory association between the backward cue of one trial and delivery of the next reward. We delivered green light (532 nm, 16–18 mW output) into the VTA 500 ms before the onset of the cue and continuing for 2 s. Rats received three reminder sessions of this training; reminder 1 occurred after instrumental conditioning, reminder 2 occurred after PIT test, and reminder 3 occurred after the incongruent/congruent test. Reminder sessions also included optogenetic inhibition identical to original training.

Instrumental training

Rats received 8 consecutive days of instrumental conditioning. Each day consisted of two training sessions separated by at least 3 hours. In each session, left or right lever was extended for 30 minutes or until 20 rewards had been received. Lever and reward relationships were fully counterbalanced as was the time of day (early vs late) for each session. Lever pressing was continuously reinforced for the first 2 days of training, reinforced on a random ratio 5 schedule for days 3-5, and reinforced on a random ratio 10 schedule for days 6-8. Rats received a reminder RR10 session in between the two PIT tests. Data are presented as total number of lever presses per session/day.

Transfer test

Rats received 2 transfer test sessions. The sessions were separated by 2 rest days and one RR10 instrumental reminder session. The data is collapsed between the two days and a 2 (Day 1 vs Day 2) x 2 (Same-Baseline vs Different-Baseline) x 2 (eYFP vs NpHR) mixed measures ANOVA revealed no significant effect of day: $F_{1, 16} = 2.373$, $p = 0.143$, no interaction between day and group: $F_{1, 16} = 0.240$, $p = 0.631$, nor interaction between day and lever: $F_{1, 16} = 0.565$, $p = 0.463$. At the start of the session, both levers were extended for 8 min to allow for extinction to the levers. All rats then received the following order of stimulus presentation: white-noise, clicker, clicker, white-noise, clicker, white-noise, white-noise, clicker, as is standard in the field.[21–23] Thus, each cue was presented 4 times for 60 s. Because cues are counterbalanced relative to the rewards they predict, the order of cue presentation is also counterbalanced in the above order. Lever pressing during the cue is subtracted from a 60 s baseline (average of lever pressing made to both levers prior to each cue presentation). This gives us a measure of how much rats increase (or decrease) responding from baseline during the cues. Data are presented as average lever presses-baseline per minute. Trials were separated by a fixed ITI of 180 s.

Forward conditioning with visual cues

Rats received 3 consecutive days of Pavlovian training where a visual cue (house light or flashing white lights) predicted the occurrence of an outcome (sucrose pellet or maltodextrin solution). Visual cues were randomly presented 15 times each for a fixed duration of 30 s and immediately terminated with the delivery of the outcome. We delivered green light (532 nm, 16–18 mW output) into the VTA

500 ms before the onset of the cue and continuing for 2 s. Responding during the visual cue is measured relative to the number of entries made 30 s before the cue conditioned stimulus (CS) was presented (i.e., CS-preCS). Data are presented as average entries per minute. Trials were separated by a variable ITI ranging from 130-230 s with an average of 180 s. Rats received two consecutive reminder sessions of this training after completing the congruency test session and before the summation test.

Congruency test

Rats received a single test session responding to congruent/incongruent audiovisual compounds presented in extinction. Four unique compounds (2 congruent and 2 incongruent) were presented four times each. Compounds were presented in the following order: clicker_flash, noise_house, noise_flash, clicker_house, noise_house, clicker_flash, clicker_house, noise_flash. Compounds were presented for a total of 30 s and were measured relative to responding made 30 s prior to compound presentation. Data are presented as average entries per minute. Trials were separated by a variable ITI ranging from 130-230 s with an average of 180 s.

Summation test

A subset of rats (N=11) received a single summation test in which the visual cues were presented by themselves or in compounds with the specific auditory cue associated with the same outcome (congruent compound). Each visual cue and audiovisual compound was presented 4 times each for a total of 16 trials. Order of presentation was pseudo-randomly counterbalanced. Cues were presented for a total of 30 s and are measured relative to responding made 30 s prior to compound presentation. Data are presented as average entries per minute. Trials were separated by a variable ITI ranging from 130-230 s with an average of 180 s.

VTA_{DA} neuronal inhibition at cue onset in forward conditioning

Rats received 8 consecutive days of Pavlovian training in a novel context where novel auditory cues (siren and pure tone) predicted the occurrence of an outcome (sucrose pellet or maltodextrin solution). Auditory cues were randomly presented 15 times each for a fixed duration of 30 s and immediately terminated with the delivery of the reward. Laser light was delivered for 2.5 s beginning 0.5 s before cue onset for one of the two cues (counterbalanced). Responding during the cues was measured relative to the number of entries made 30 s before the cue was presented (CS-preCS). Trials were separated by a variable ITI ranging from 130-230 s with an average of 180 s. After 8 days of conditioning, rats received a single test session in extinction where each stimulus was presented 8 times without laser delivery. Stimulus presentation was pseudo-randomly ordered and fully counterbalanced. Auditory cues were presented for a total of 30 s and are measured relative to responding made 30 s prior to cue presentation. Trials were separated by a variable ITI ranging from 130-230 s with an average of 180 s. Data are presented as average entries per minute.

Histology

The rats were euthanized with an overdose of carbon dioxide and perfused with phosphate-buffered saline followed by 4% paraformaldehyde (Santa Cruz Biotechnology). Fixed brains were cut in 20-μm sections, and images of these brain slices were acquired and examined under a fluorescence microscope (Carl Zeiss Microscopy). The viral spread and optical fiber placement (Figures 2A and 2B) were verified and later analyzed and graphed using Adobe Photoshop. Any rats with viral expression or placement outside of the VTA were removed from all analyses.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data was collected using Med-Associates automated software and the text file output were analyzed using MPC2XL (Med Associates, St. Albans, VT). Repeated Measures Analysis of Variance (ANOVA) were used to assess training and test data in JASP (version 0.15). Simple effects were used to follow up on significant interactions and assess the effect of lever (Same vs Diff) on each group (eYFP vs NpHR), the effect of compound type (Incongruent vs Congruent) on each group, and the effect of cue type (Visual CS+ vs Compound) on each group. Statistical significance was defined as a p value less than 0.05. One sample t tests were used to measure responding relative to baseline (expected value = 0). All data were tested for normality and analyses that did not pass this criterion were adjusted using a Greenhouse-Geisser (Repeated Measures) or Wilcoxon (t test) correction. For instances in which a Greenhouse-Geisser correction was used, the adjusted p value is reported but degrees of freedom are reported in their uncorrected form. Pilot data (n=11) presented in the supplementary material revealed the effect of lever on the PIT test was very large, $\eta^2 = 0.519$ or $f = 1.039$ using the formula ($f = \mathrm{sqr}(\eta^2 / (1 - \eta^2))$). A power analysis conducted in G*power (version 3.1) revealed 8 participants would be necessary to discover a similarly sized effect with 90% power (between measurement $r = 0.074$). Thus, we were well powered to detect a main effect of lever in our initial PIT test with 9 participants per group. All details regarding sample sizes can be found in the results section of the manuscript.