Article

Current Biology

Dopamine release in the nucleus accumbens core signals perceived saliency

Highlights

- NAc core dopamine only mimics reward prediction error in select reward contexts
- RPE does not model dopamine release during negative reinforcement
- Dopamine signaling in the NAc core does not support valence-free prediction error
- NAc core dopamine tracks valence-free perceived saliency in all conditions

Authors

Munir Gunes Kutlu, Jennifer E. Zachry, Patrick R. Melugin, ..., Lin Tian, Cody A. Siciliano, Erin S. Calipari

Correspondence

erin.calipari@vanderbilt.edu

In brief

Kutlu et al. examine dopaminergic information encoding in the nucleus accumbens. This work revises the long-held theory of dopamine as a reward prediction molecule and provides a novel framework for dopamine as a perceived saliency signal. These results unite multiple theories of dopamine signaling and have broad implications for psychopathologies.

Kutlu et al., 2021, Current Biology *31*, 4748–4761
November 8, 2021
https://doi.org/10.1016/j.cub.2021.08.052

CellPress

CellPress

Current Biology

Article

Dopamine release in the nucleus accumbens core signals perceived saliency

Munir Gunes Kutlu,[1] Jennifer E. Zachry,[1] Patrick R. Melugin,[2] Stephanie A. Cajigas,[1,2] Maxime F. Chevee,[1] Shannon J. Kelly,[1] Banu Kutlu,[1,7] Lin Tian,[3] Cody A. Siciliano,[1,2,4] and Erin S. Calipari[1,2,4,5,6,8,9,*]

[1]Department of Pharmacology, Vanderbilt University, Nashville, TN 37232, USA
[2]Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN 37232, USA
[3]Department of Biochemistry and Molecular Medicine, University of California, Davis, Sacramento, CA 95817, USA
[4]Vanderbilt Center for Addiction Research, Vanderbilt University, Nashville, TN 37232, USA
[5]Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN 37232, USA
[6]Department of Psychiatry and Behavioral Sciences, Vanderbilt University, Nashville, TN 37232, USA
[7]Libraries Strategic Technologies, Penn State University Libraries, University Park, PA 16802, USA
[8]Twitter: @TheCaliparilab
[9]Lead contact
*Correspondence: erin.calipari@vanderbilt.edu
https://doi.org/10.1016/j.cub.2021.08.052

SUMMARY

A large body of work has aimed to define the precise information encoded by dopaminergic projections innervating the nucleus accumbens (NAc). Prevailing models are based on reward prediction error (RPE) theory, in which dopamine updates associations between rewards and predictive cues by encoding perceived errors between predictions and outcomes. However, RPE cannot describe multiple phenomena to which dopamine is inextricably linked, such as behavior driven by aversive and neutral stimuli. We combined a series of behavioral tasks with direct, subsecond dopamine monitoring in the NAc of mice, machine learning, computational modeling, and optogenetic manipulations to describe behavior and related dopamine release patterns across multiple contingencies reinforced by differentially valenced outcomes. We show that dopamine release only conforms to RPE predictions in a subset of learning scenarios but fits valence-independent perceived saliency encoding across conditions. Here, we provide an extended, comprehensive framework for accumbal dopamine release in behavioral control.

INTRODUCTION

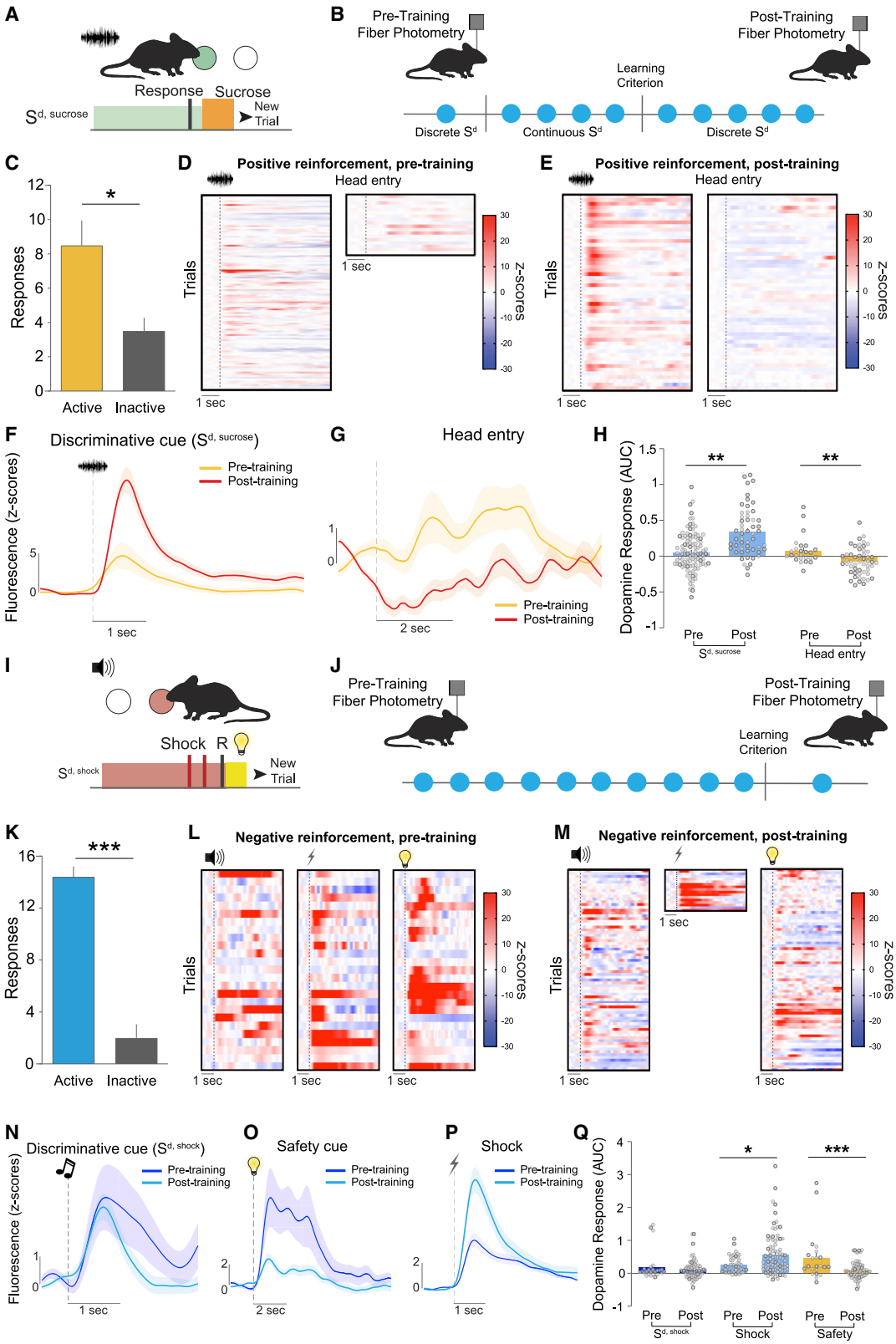
There has been a great deal of work aimed at understanding the role of dopamine in learning and memory.[1–8] The prevailing theory is that dopamine neurons projecting to the ventral striatum are the biological substrate for reward prediction error (RPE) and they transmit information about rewards and predictive cues and update this information when errors in predictions are encountered.[9,10] The prediction error hypothesis of dopamine signaling originates from an influential Pavlovian conditioning model, the Rescorla-Wagner model,[11] which assumes that learning occurs when outcomes are not perfectly predicted. Biological evidence for dopamine neurons encoding an RPE signal was first demonstrated by Schultz and colleagues,[1] showing increases in dopamine neuron firing rates when an unexpected reward is encountered, a learning-dependent shift in firing to cues that predict reward delivery, and a decrease in firing rates when expected rewards are withheld. Similar outcomes have been observed across species in ventral tegmental area (VTA) cell bodies[3,12–15] and in dopamine release in the nucleus accumbens (NAc).[4,16–21] As a result, the RPE model of the role of dopamine in learning has been dominant in the field for the last 20 years. Nevertheless, there is literature suggesting that the role of dopamine in learning and memory deviates from RPE.[18,22–32] Furthermore, the RPE hypothesis has typically been tested under limited behavioral contingencies, largely relegated to reward-based contexts and in more limited contexts with aversive stimuli.

Here, we investigate the role of dopamine in the NAc core across different types of learning paradigms to understand how dopamine signaling maps onto standing theories of learning and memory. We show that dopamine contributes to learning about diverse contingencies by signaling the perceived saliency of stimuli. These results unite multiple theories of dopamine signaling and have broad implications for our understanding of behavioral control and neuropsychiatric disorders.

RESULTS

Dopamine release does not track RPE in aversive contexts

To disentangle multiple task parameters (e.g., valence, action initiation, prediction[33]), we used the recently developed multidimensional cue outcome action task (MCOAT)[34] in combination with the genetically encoded dopamine sensor, dLight1.1,[35] to record *in vivo* dopamine dynamics in the NAc core (Figures S1A–S1D). We trained mice in a positive reinforcement task, in which an auditory cue predicted that an operant response would result in the delivery of sucrose (positive



(legend on next page)

reinforcement) (Figures 1A and 1B). Subsequently, mice were trained in an aversive learning task, in which a distinct auditory cue indicated that a different response would prevent the delivery of footshocks (negative reinforcement) (Figures 1I and 1J). The operant response in the two phases requires the same motoric action (nose poke), and correct responses produce outcomes with positive valence (sucrose retrieval/shock removal); however, the reinforcer has opposite valence (sucrose, positive; shock, negative). If dopamine is critical in the encoding of positive outcomes, then the neural signatures should be similar when positive and negative reinforcement tasks are successfully performed.

We recorded NAc core dopamine responses to an auditory discriminative cue ($S^{d,\ sucrose}$) and the contingent delivery of sucrose (positive reinforcement; Figures 1A–1C, S1E, and S1F). Dopamine responses to the $S^{d,\ sucrose}$ increased from pre- to post-training, whereas dopamine responses during sucrose retrieval decreased (Figures 1D–1H and S1K). In agreement with the previous literature,[1,15,16,19] this pattern follows what is suggested by the Rescorla-Wagner model and RPE-based accounts of dopamine in reward-based predictions (Figures S1G and S1H).

During negative reinforcement, a separate auditory cue ($S^{d,\ shock}$) signaled the ability to nose poke to avoid a series of footshocks (Figures 1I and 1J). Mice showed robust negative reinforcement learning (Figures 1K, S1I, and S1J). In contrast to RPE predictions, dopamine response to the $S^{d,\ shock}$ did not change during negative reinforcement learning, even though the cue became predictive of the avoidance of footshock (Figures 1L–1N and 1Q). Furthermore, dopamine was evoked by the safety cue; however, the response was largest when the safety cue was novel and before the valence could have been attributed (Figures 1O, 1Q, and S1L). This is surprising, as RPE accounts considered dopamine responses to safety cues as a part of reward processing.[36,37] However, previous studies primarily analyzed dopamine responses to safety cues after they gained positive valence, rather than over the entire learning process. It is important to note that the safety cue itself is not inherently rewarding and it only acquires value after the animal learns that it predicts the removal/avoidance of an aversive event.[38] Lastly, dopamine responses to the aversive outcome (footshock) were positive and increased with training (Figures 1L, 1M, 1P, and 1Q), which was not due to movement (Figures S2A–S2G).

Dopamine responses to aversive outcomes, not safety cues, predict future behavior

The results showing that the dopamine signal to the safety cue goes down over learning have been interpreted as being in line with RPE models.[36,37,39–41] While the data presented suggest that this is not the case, if the RPE interpretation is correct, then dopamine to the safety cue would serve as an error signal to update subsequent decisions. To test this, we used a supervised machine learning approach to examine whether features of the dopamine signal are predictive of behavioral performance on a trial-by-trial basis. We used a support vector machine (SVM) in which we iteratively divided our data into training and test sets to determine whether we could accurately predict each mouse's behavior based on the dopamine response to various stimuli in each trial (Figure 2A).

The SVM was able to predict whether a response was made in the current trial based on the dopamine response to the $S^{d,\ sucrose}$ signaling positive reinforcement (Figures 2B–2E), but not to the $S^{d,\ shock}$ signaling negative reinforcement (Figures 2F–2I and S3). The dopamine signal that occurred following shock delivery was able to successfully predict the next trial behavioral response (whether the animal would respond during the $S^{d,\ shock}$ on the next trial; Figures 2J and S3G–S3K), suggesting that this dopamine response played a critical role in driving future behavior, even though the outcome itself was aversive and the dopamine response was positive.

The SVM was unable to use the dopamine response to the safety cue to predict the behavioral response on the subsequent trial above chance (Figures 2K, S3L, and S3M). Importantly, after training, dopamine responses to safety cues were larger in trials in which a change in future behavior was not necessary compared to those in which updating would improve performance (Figures S3N–S3P). Therefore, contrary to the interpretation that the dopamine response to the safety cue serves as a positive outcome and thus an error-based updating signal, this signal does not track errors needed to update future behavior.

Accumbal dopamine tracks stimulus saliency

While these data show that dopamine does not signal RPE, dopamine release could still signal other necessary components of associative learning, such as saliency, as others have suggested.[42,43] The physical intensity of a stimulus (e.g., amperage, decibel, lux), called saliency, is a principal component of learning.[11] To probe how dopamine release changes to scaling

Figure 1. Dopamine in the nucleus accumbens core does not track reward prediction error in negative reinforcement tasks

(A) Positive reinforcement task.

(B) Dopamine was recorded via fiber photometry using dLight1.1.

(C) Mice made more active than inactive responses during the post-training session ($t_6 = 3.18$, $p = 0.024$; $n = 6$).

(D and E) Heatmaps showing dopamine responses aligned around $S^{d,sucrose}$ (left) and head entry (right) during (D) pre-training and (E) post-training. Each row represents a single $S^{d,sucrose}$ presentation or head entry.

(F and G) Averaged traces showing $S^{d,sucrose}$ (F) and (G) head entry responses during pre- and post-training.

(H) Dopamine increased to the $S^{d,sucrose}$ ($F_{1,153} = 10.79$, $p = 0.0013$) and decreased to head entries over training ($F_{1,75} = 11.17$, $p = 0.0013$).

(I) Negative reinforcement task.

(J) Fiber photometry recording design.

(K) Behavioral performance during negative reinforcement post-training session ($t_4 = 9.35$, $p < 0.001$; $n = 5$).

(L and M) Heatmaps of dopamine responses aligned around $S^{d,shock}$ (left), footshock (center), and safety cue (right) during pre-training (L) and (M) post-training.

(N–P) Dopamine traces to $S^{d,shock}$ (N), (O) safety cue, and (P) footshock pre- and post-training.

(Q) Dopamine response to $S^{d,shock}$ ($F_{1,96} = 1.52$, $p = 0.220$), footshock ($F_{1,127} = 4.00$, $p = 0.047$), and safety cues ($F_{1,95} = 15.46$, $p = 0.0002$). Error bars represent ± S.E.M. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

See Figures S1 and S2.

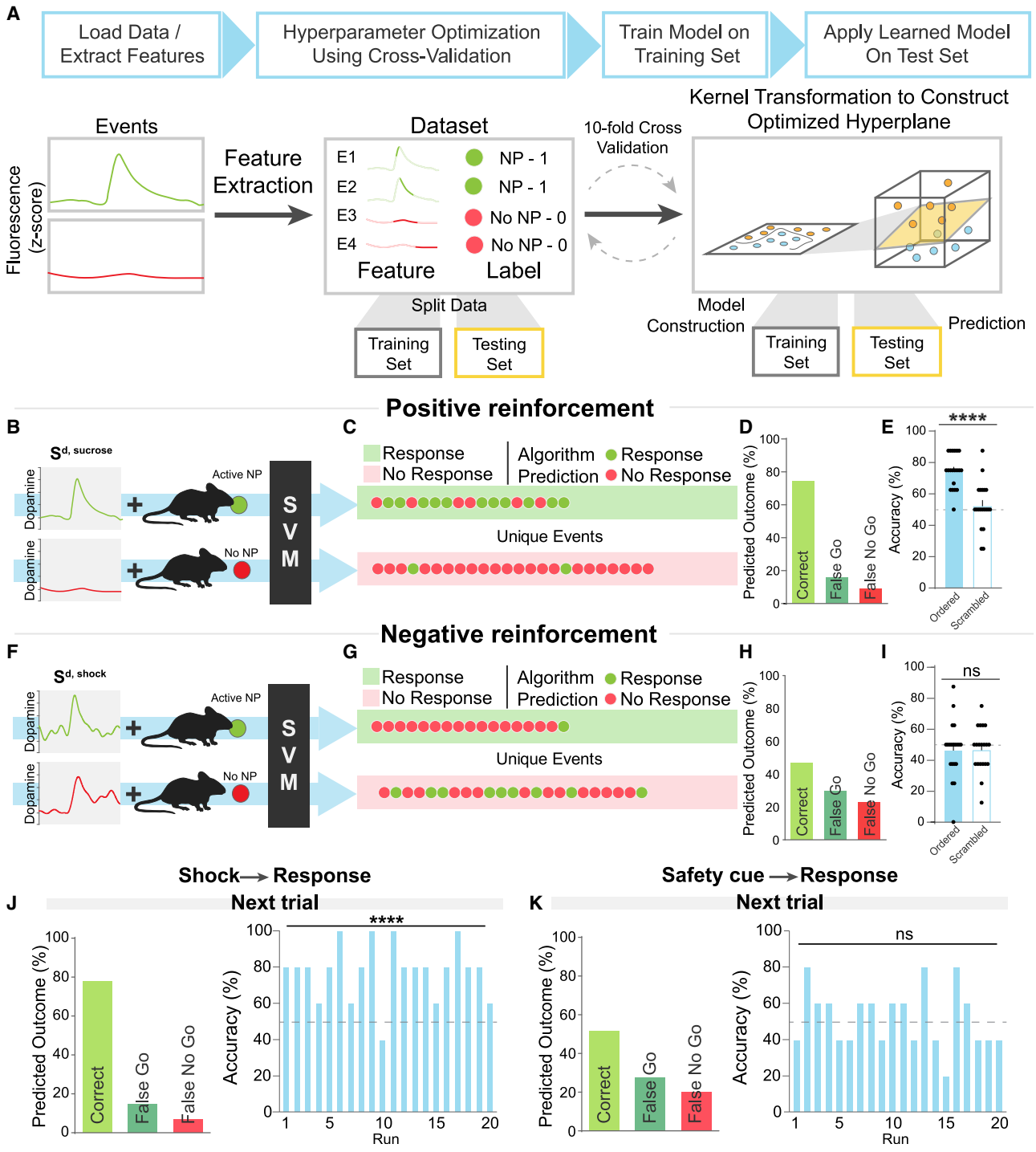


Figure 2. Dopamine responses to footshocks, but not safety cues, predict future behavior

(A) Support vector machines (SVM) were used to define whether dopamine predicted current and/or future behavior.

(B–E) The SVM accurately predicted current trial behavior based on dopamine responses to the $S^{d,sucrose}$ ($t_{38} = 5.34$, $p < 0.0001$; N = 69 trials) with few errors (opposite dot color).

(F–I) The SVM was unable to predict current trial behavior based on dopamine responses to the $S^{d,shock}$ ($t_{38} = 0.000$, $p > 0.99$; N = 69 trials).

(J) The SVM was able to predict whether an animal would respond during the $S^{d,shock}$ on the next trial based on the dopamine response to the shock itself ($t_{19} = 7.95$, $p < 0.0001$; N = 30 trials; see also Figure S3).

(K) The SVM was unable to predict future trial responses based on the dopamine response to the safety cue ($t_{19} = 0.55$, $p = 0.59$; N = 30 trials; see Figure S3). Error bars represent ± S.E.M. **** $p < 0.0001$; ns, not significant.

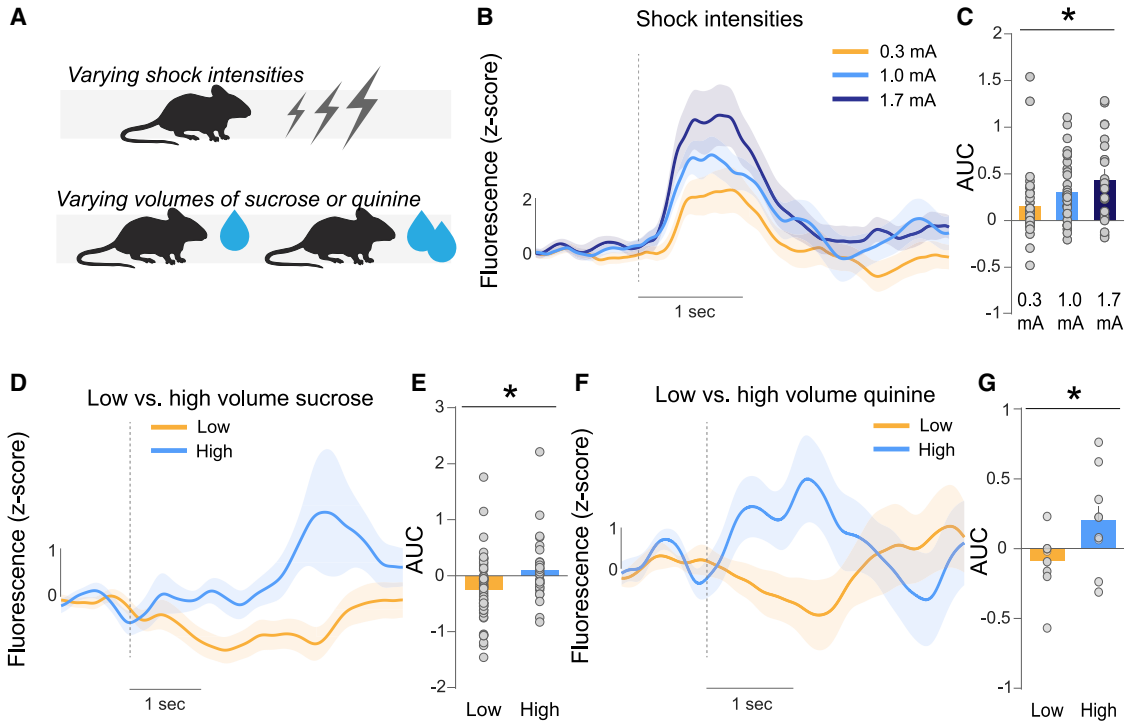


Figure 3. Dopamine release tracks saliency

(A) Footshock intensity, sucrose volume, or quinine volume was manipulated.

(B–G) Dopamine response in the NAc core increased with shock intensity ($F_{2,83} = 4.16$, $p = 0.019$; 0.3 mA versus 1.7 mA $p = 0.008$; $n = 5$), (D and E) increased with larger sucrose volumes ($F_{1,85} = 6.31$, $p = 0.014$; $n = 3$), and (F and G) increased with larger quinine volumes ($F_{1,18} = 4.60$, $p = 0.048$; $n = 3$). Error bars represent ± S.E.M. * $p < 0.05$.

stimulus intensity, we recorded dopamine to positive and negative stimuli of varying intensities (Figure 3A). Dopamine responses to footshocks increased as intensity was increased (Figures 3B and 3C). Next, mice were trained to nose poke for sucrose on a continuous fixed-ratio one schedule of reinforcement, without the requirement of a discriminated operant response. The volume of sucrose was increased on a random subset of trials (Figure 3A). Dopamine release increased with sucrose volume (Figures 3D and 3E). Finally, animals were exposed to the bitter tastant quinine. Dopamine responses were increased to higher volumes of quinine (Figures 3F and 3G). These results show that NAc core dopamine is influenced by stimulus intensity in a fashion that is independent of stimulus valence.

Dopamine release does not track prediction error hypotheses during unexpected presentation or omission of stimuli

While the previously presented data suggest a strong role for encoding novel and salient events, it was still possible that dopamine signals prediction error for valenced stimuli without reflecting valence directionality (i.e., unsigned prediction error). Mice were trained in a single fear conditioning session, in which they received 11 tone-footshock pairings (Figure S4A). During this session, tone was followed by shock for 100% of trials. Aversive footshocks resulted in a robust positive dopamine response (Figures 4A, 4B, S4A, and S4B), and this signal stayed positive throughout the fear conditioning session (Figures S4A–S4C).

After initial training, shock was omitted in 20% of trials (Figures 4A–4E). There were no differences in freezing (Figure 4A) or dopamine responses to the cues (Figure S4D and S4E) during omission trials. The omission of the footshock resulted in a positive dopamine response at the time of the expected footshock, although the omission response was smaller than when the shock was present (Figures 4B and 4C). Therefore, NAc core dopamine is increased when an expected prediction is not met, rather than a signed prediction error (which would be positive during the addition of an unexpected stimulus and negative during the omission of an expected stimulus). However, this also ruled out an unsigned prediction error, as the dopamine response to the omitted stimulus was smaller than when the stimulus was present, rather than larger to signal a deviation from prediction. We also found that the dopamine response to the footshock was stronger during the omission session, in which the footshocks were presented on only 80% of the trials compared to the first fear conditioning session, in which shocks were presented with 100% probability (Figures 4D and 4E), likely because the outcome is less certain and thus perceived as more salient when it does occur.

The dopamine release pattern during extinction further ruled out both signed and unsigned prediction error accounts. Freezing to shock-predictive cues progressively decreased with extinction (Figure 4F); however, the dopamine response to these cues did not change (Figures 4G and 4H), demonstrating, along with additional studies with appetitive cues (Figures S4F–

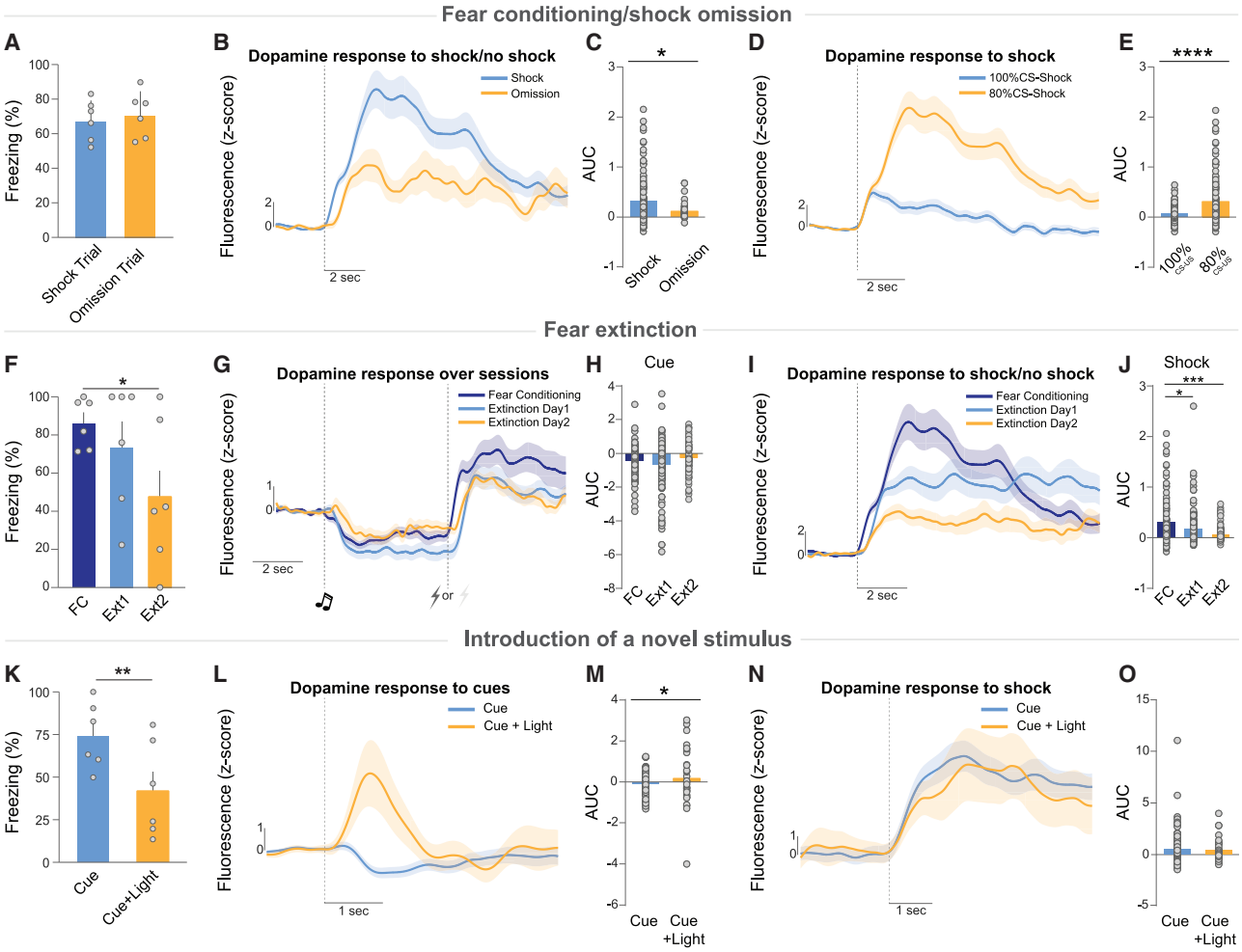


Figure 4. Dopamine release does not track prediction error hypotheses

(A) After fear conditioning training, footshocks were omitted in 20% of trials. Freezing during each trial type was not different ($t_5 = 1.47$, $p = 0.20$; $n = 6$).

(B and C) Omitted footshocks evoked a dopamine transient ($t_{29} = 3.92$, $p = 0.0005$), which was lower in amplitude than when the footshock was present ($F_{1,156} = 7.39$ $p = 0.0073$).

(D and E) Dopamine responses to footshocks were larger when cue-shock probability was 80% as compared to sessions in which cue-shock probability was 100% ($F_{1,222} = 21.48$, $p < 0.0001$).

(F) Freezing during acquisition and extinction ($t_5 = 3.31$, $p = 0.021$; $n = 6$).

(G and H) Dopamine response to the cue was not changed by extinction ($F_{2,331} = 3.37$, $p = 0.035$; fear conditioning versus early and late extinction, $p > 0.05$).

(I and J) Dopamine responses to the shock delivery period decreased over extinction ($F_{2,329} = 9.76$, $p = 0.0001$; fear conditioning versus early extinction; $p = 0.0005$; fear conditioning versus late extinction; $p < 0.0001$).

(K) Introduction of a novel cue reduced freezing ($t_5 = 4.23$, $p = 0.008$; $n = 6$).

(L and M) Dopamine response to the tone + novel cue was increased ($F_{1,156} = 6.44$, $p = 0.012$).

(N and O) Dopamine responses to footshock did not differ between trial types ($F_{1,156} = 0.35$, $p = 0.56$).

See Figure S4. Error bars represent ± S.E.M. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

S4K), that cue responses are not a function of learned negative valence (see Figures S4L–S4Q). In addition, NAc core dopamine responses to footshocks were strongest during fear conditioning—when shock was present—and progressively decreased over extinction (Figures 4I and 4J).

To parse whether (1) dopamine signals the associative strength of cues and (2) increases in dopamine release add value to environmental stimuli/cues, we introduced a novel stimulus after initial conditioning (Figures 4K–4O). The novel cue (light) was presented concurrent with the original predictive cue (tone) and was followed by a footshock. Learning theory and empirical results dictate that because the conditioned response to the tone has already been acquired, the cue adds no new information and, as such, does not acquire value. However, novel stimuli can decrease conditioned responses when presented together with a conditioned stimulus, a phenomenon known as external inhibition.[44]

The novel stimulus reduced freezing (Figure 4K), demonstrating external inhibition. Although there was no error in prediction, a robust positive dopamine transient was observed to the light + tone presentation, but not to the tone alone (Figures 4L and 4M). Furthermore, the dopamine response to the footshocks

did not differ across trial types (Figures 4N and 4**O**). Thus, dopamine responds to valence-free stimuli that do not signal a change in associative strength.

Dopamine responds to stimulus saliency and is dynamically modulated by novelty

These data show that physical intensity determines dopamine responses to stimuli, but this alone cannot account for learning-induced changes in dopamine dynamics. Importantly, physical intensity is perceived subjectively in different contexts and situations, a construct known as perceived saliency.[45–47] Perceived saliency contributes to learning and can change over experience; therefore, it could provide a parsimonious explanation to seemingly disparate results showing that dopamine signals physical stimulus saliency but also is modulated across learning.[4,16]

Perceived saliency is highly influenced by the novelty of the environment/stimulus. Novelty is high when a stimulus is encountered for the first time and decreases with repeated exposure.[48–52] If NAc core dopamine tracks perceived saliency, then it should initially track physical intensity (see Figure 3), but also be dynamically influenced by novelty (as observed in Figures 4K–4O); thus, we hypothesized that dopamine release would decrease with the repeated presentation of a stimulus despite the valence and intensity being held constant.

We presented repeated footshocks with the same physical intensity at fixed time intervals (Figure 5A). Supporting our hypothesis, the amplitude of dopamine release to the footshocks decreased with repeated presentations (Figures 5A and 5B). This was not due to the detection limits of shock-evoked dopamine release (Figures S5A–S5E) or changes in baseline fluorescence and was insensitive to the methodology for quantifying transient magnitude (Figures S5F–S5J). This pattern was observed whether the footshocks were presented in an unsignaled fashion (Figures 5A and 5B) or in a negative reinforcement context (Figures S5K–S5Q). Therefore, as footshocks become less novel, thereby decreasing perceived saliency, the magnitude of the dopamine signal decreases.

We next sought to parse whether perceived saliency encoding could be observed in scenarios that can be fully described by RPE. Following acquisition of positive reinforcement (Figure 5C), the outcome was switched so that a nose poke in the presence of the discriminative stimulus (S^d) resulted in a footshock (no sucrose delivery), while all other task parameters remained the same (Figure 5E). This contingency switch explicitly produces an error in a reward-based prediction; thus, if dopamine signals RPE, the dopamine response should be negative during the first footshock and responses to the S^d should decrease as the value is updated. However, if dopamine signals saliency or perceived saliency, a positive response would be expected given that the footshock has never been paired previously with the operant response and is therefore both novel and salient. If dopamine signals perceived saliency, but not saliency per se, then the response to the predictive cue should also increase as the familiar cue acquires a novel prediction.

Mice decreased and eventually stopped responding (Figure 5D), demonstrating that the shock functioned as a punisher and progressively updated the conditioned association of the S^d. Confirming the interpretation that dopamine responds to novel and salient events, the first footshock following the contingency switch elicited a positive dopamine transient (Figure 5F). After just a single trial, the dopamine response to the S^d following the unexpected shock was increased, even though the cue was familiar, presented at a fixed intensity, and now predicted a negative outcome (Figures 5G and 5H). This demonstrates that dopamine is increased when a novel stimulus is introduced and is increased by familiar stimuli when they convey novel information, even when there is no change in physical saliency or familiarity of the stimulus itself.

If perceived saliency is the underlying construct that can explain NAc dopamine responses across conditions, dopamine should also respond to stimuli even in the absence of acquired or innate valence. Neutral stimuli evoked dopamine release (Figures 6A–6C), which decreased over repeated exposure (Figures 6D–6F), as would be predicted for a perceived saliency signal.

Next, we repeatedly presented a mild footshock (0.3 mA, 50 presentations). The dopamine response habituated with repeated presentations (Figures 6G and 6H), and the additional trials in this experiment revealed that the dopamine response decreased to baseline (Figures 6H and 6I), again demonstrating that physical intensity alone cannot explain dopamine responses. On the 51st presentation, we increased the intensity of the footshock to 1 mA, which had not been experienced before and is therefore both salient and novel. The dopamine response to shock was again observed, despite being habituated to the 0.3-mA shock just prior (Figures 6I and 6J). Although the physical intensity was now stronger, the response was not larger than the response to the initial 0.3-mA shock (Figures 6K and 6L). Thus, perceived saliency, which is a product of both novelty and physical intensity, can explain NAc dopamine release patterns.

A novel model of behavioral control: the Kutlu-Calipari-Schmajuk (KCS) model

These data show that theories used to explain dopaminergic information encoding do not hold up as predictive models of dopaminergic responses when they are pushed outside the narrow parameters they were originally designed to explain. Even more recent accounts of Pavlovian conditioning[32,53] fall short, as they cannot explain concepts like latent inhibition and sensory preconditioning which have been shown to be directly altered by dopamine manipulations.[54,55] To address this, we developed a new behavioral model, the KCS model - based on an earlier neural network model of Pavlovian conditioning (the SLGK model [56]) -, which allows unbiased mapping of dopamine onto its theoretical components (i.e., prediction error, association formation, attention, and temporal dynamics).

At the core of the model (Figure 7A; see Method details for a complete list of equations) is an error prediction term where associations are formed based on Rescorla-Wagner-based predictions. An additional aspect of this model is perceived saliency, which is computationally defined as the product of stimulus intensity and the attentional value of a stimulus. The core factor that controls attentional allocation is the level of novelty in each context. Accordingly, perceived saliency increases when novelty is high, and the organism directs more attention to that stimulus even when the physical intensity is constant.

The KCS model can replicate basic learning phenomena that the Rescorla-Wagner model was developed to describe,[56] such as blocking[57] and overshadowing[58] (Figure S6A), as well

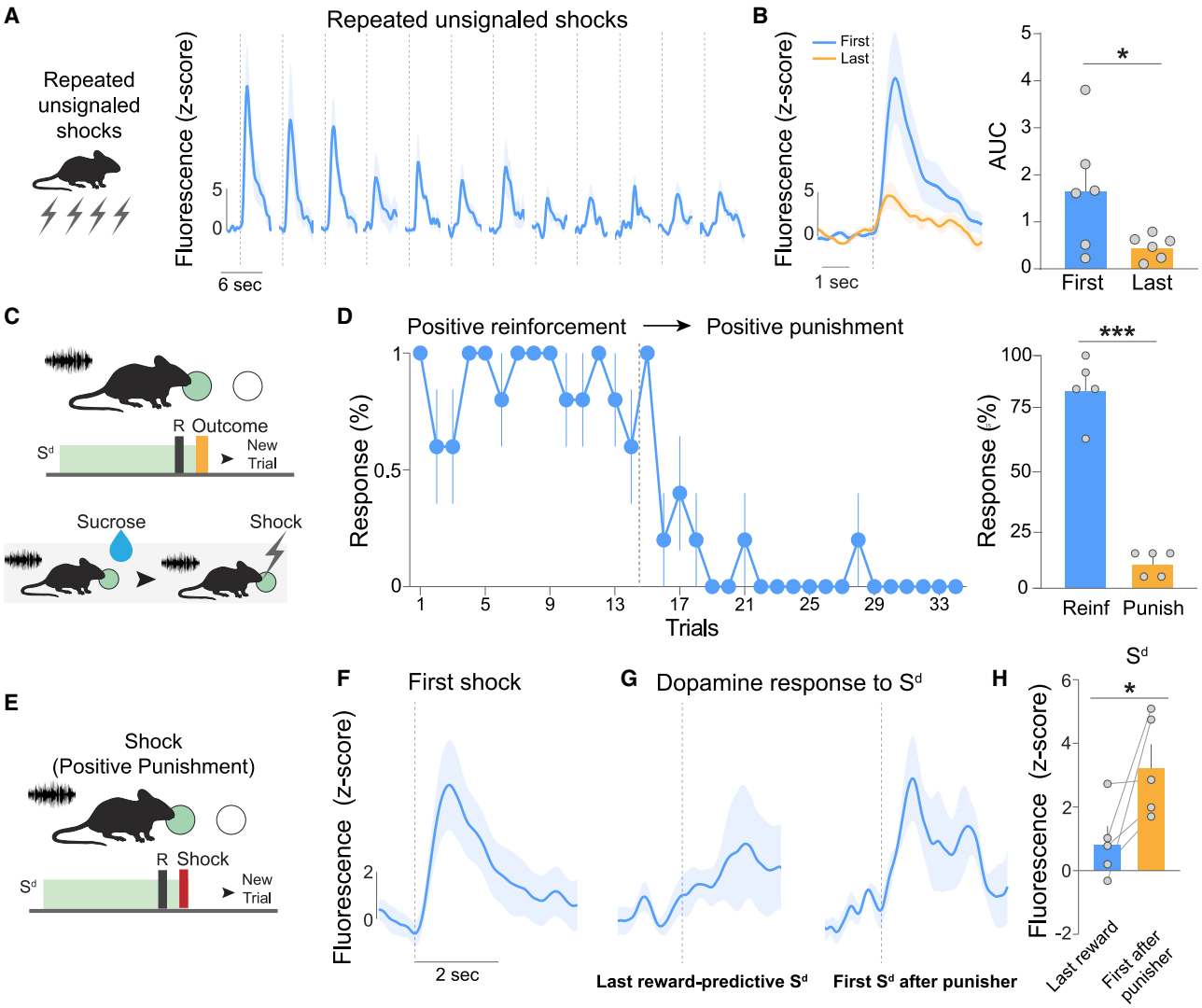


Figure 5. Dopamine decreases with repeated presentations of stimuli, increases during worse than expected outcomes

(A) Mice received repeated footshocks with constant intensity.

(B) Dopamine responses to footshocks decreased over presentations ($t_5 = 2.60$, $p = 0.047$; $n = 6$).

(C) Mice nose poked during a discriminative cue (S^d) for sucrose. In subsequent trials, responses during the same S^d now resulted in a footshock.

(D) Switching from positive reinforcement to punishment (denoted by dotted line) decreased operant responses ($t_4 = 11.57$, $p = 0.0003$; $n = 5$. Reinf: positive reinforcement, Punish: positive punishment).

(E) During the contingency switch, the S^d resulted in footshock delivery and represented a worse than expected outcome.

(F) The dopamine response to the first footshock following the operant response was positive.

(G and H) Dopamine response to the S^d increased ($t_4 = 2.55$, $p = 0.031$).

See Figure S5. Error bars represent ± S.E.M. * $p < 0.05$, *** $p < 0.001$.

as many that it cannot[50,59] (Figures S6B and S6D–S6F). It can also describe temporal learning phenomena[53] (Figure S6C) and basic operant conditioning schedules (Figures S6D–S6F).

Dopamine release in the NAc core tracks perceived saliency

Using the KCS model, we computed the predicted behavioral responses, perceived saliency, associative strength, and prediction error values for the fear conditioning and unconditioned stimulus experiments described above (Figures 4, S6, and S7). The KCS model successfully predicted behavioral response patterns during fear conditioning, footshock omission, fear extinction, and introduction of a novel stimulus (Figures S7A–S7F). The perceived saliency component of the KCS model alone tracked dopamine patterns in all cases (Figures S6G–S6M and S7G–S7L).

We simulated the experiment from Figures 5C–5H in which a contingency switch explicitly produces an error in a reward-based prediction. KCS model simulations showed that the model can predict behavioral responses (Figure 7B) and dopamine patterns (Figures 7C and 7D) that follow changes in perceived saliency but are opposite from prediction error simulations (Figures

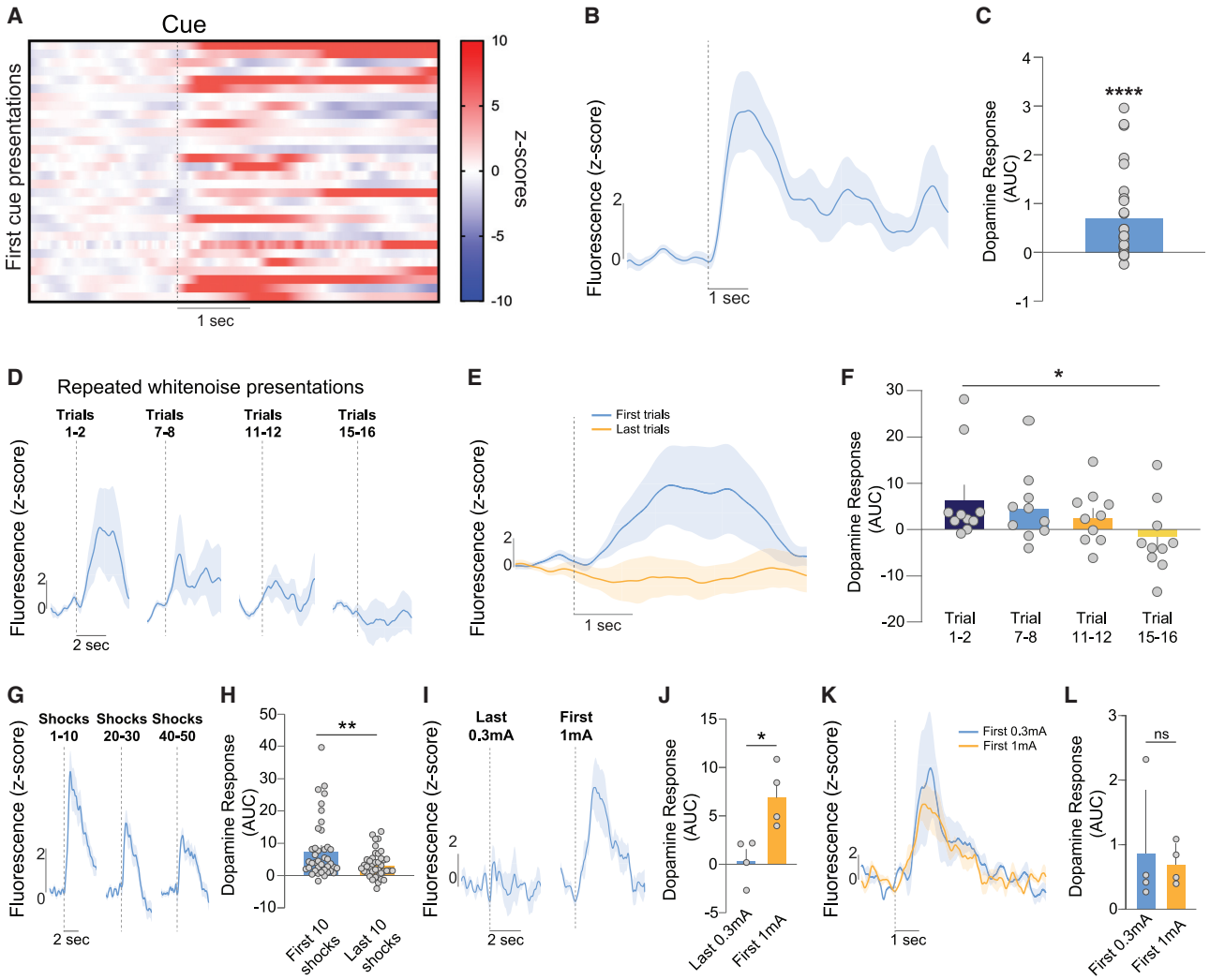


Figure 6. Neutral stimuli evoke dopamine release

(A) Heatmap showing dopamine responses to the first-time presentation of cues.

(B and C) Dopamine response to discriminative stimuli before acquiring value (t_{25} = 4.08, p = 0.0004; n = 10).

(D) Neutral white noise evoked a positive dopamine response.

(E and F) Dopamine response to white noise decreased with repeated presentation ($F_{2.594,23.35}$ = 3.20, p = 0.048; trial1–2 versus trial 15–16, p = 0.036; n = 5).

(G and H) Dopamine response to a mild footshock decreased as novelty decreased (t_{39} = 3.46, p = 0.0013).

(I and J) Increasing footshock intensity reinstated the habituated dopamine response (t_3 = 4.94, p = 0.0159).

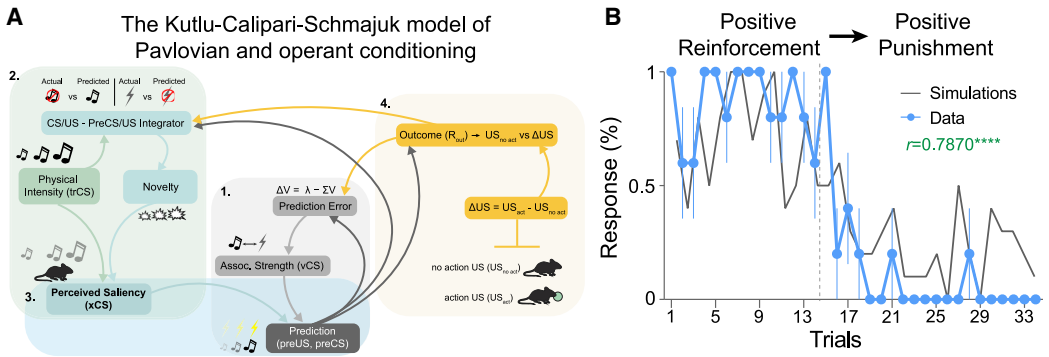
(K and L) The high intensity footshock did not yield a stronger response than the first footshock presentation at the lower intensity (t_3 = 0.47, p = 0.67; n = 4). Error bars represent ± S.E.M. * $p < 0.05$, ** $p < 0.05$, **** p < 0.0001, ns = not significant.

7E and 7F). These data show the dopamine maps onto perceived saliency across the conditions tested within this study.

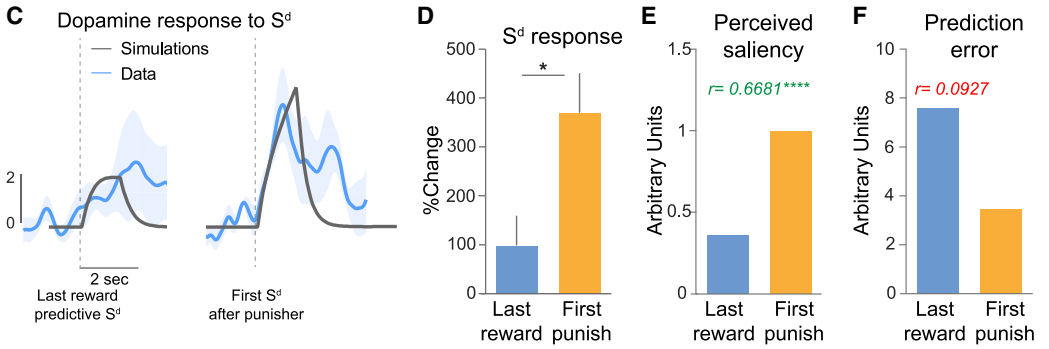
Testing model predictions via optogenetic manipulations

To stimulate dopamine terminals in the NAc core, we used an intersectional approach to achieve dopamine-specific expression of the excitatory opsin channelrhodopsin (ChR2; or EYFP for controls[60]) in the VTA (Figure 7G). By implanting a fiber optic over the NAc core, we selectively stimulated NAc core dopamine terminals (Figure S6N).

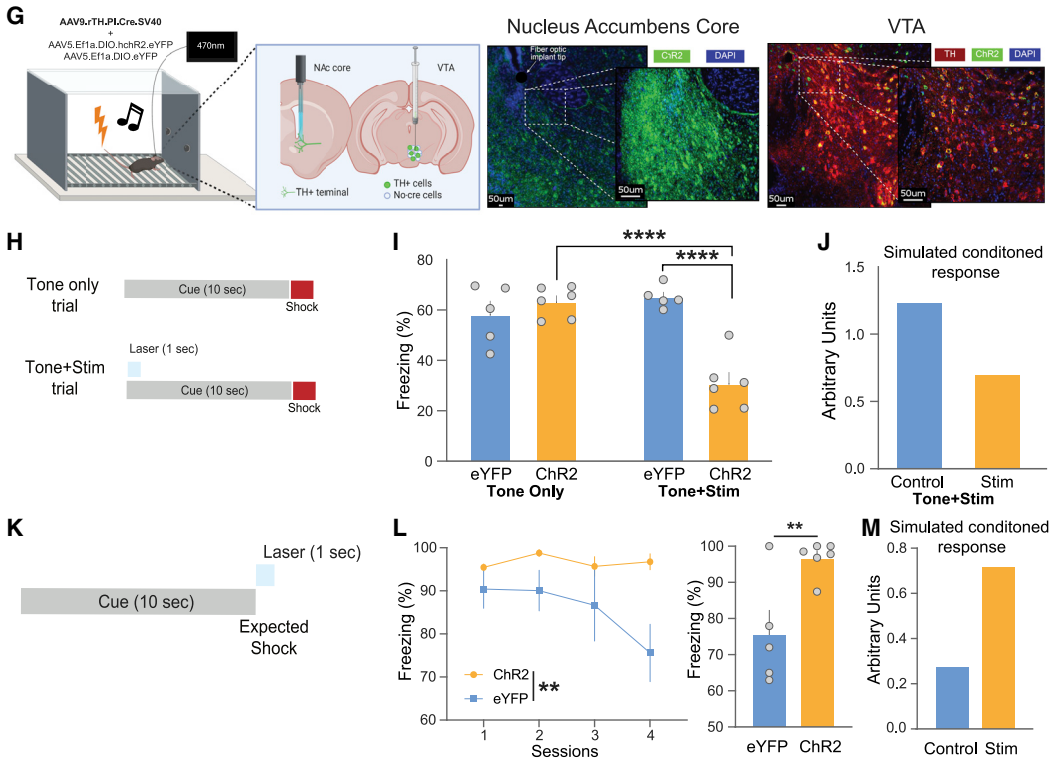
In Figures 4K–4M, we observed that novel cues were capable of both increasing dopamine levels and reducing freezing behavior (i.e., inducing external inhibition), even though there was no prediction error. If dopamine signals perceived saliency, then increasing dopamine to a previously learned fear cue would induce external inhibition and decrease freezing during those trials. Alternatively, if dopamine signals associative strength or prediction error, then increasing dopamine to the cue should increase the conditioned response. Following an initial fear conditioning training session, we stimulated dopamine terminals during 25% of cue presentations (Figure 7H). When dopamine was artificially increased during the cue, freezing was decreased (Figure 7I). When perceived saliency values for fear conditioning cues were computationally increased using the KCS model, the predicted conditioned response was similarly decreased



Dopamine tracks perceived saliency



Stimulation of accumbal dopamine enhances perceived saliency



(legend on next page)

(Figure 7J), thus confirming that dopamine transmits a perceived saliency signal.

In addition, we observed a dopamine response at the time of a predicted but absent footshock in Figures 4A–4C. If dopamine signals perceived saliency, then optogenetically enhancing it should prevent extinction (Figure 7K). Alternatively, if dopamine is transmitting an error signal, then increasing dopamine at this time would signal a larger error and thus enhance extinction. Stimulating dopamine terminals at the time of the omitted shock prevented extinction (Figure 7L). Computationally amplifying the perceived saliency of the omitted outcomes within the KCS model also similarly interfered with extinction learning (Figure 7M). These data confirm that NAc core dopamine signals perceived saliency.

DISCUSSION

Here, we provide biological experiments and an empirically verified computational model that shows that mesolimbic dopamine release patterns conform to a perceived saliency signal. These results pose significant challenges to the hypothesis of dopamine as a prediction error signal while offering an alternative account for the role of NAc core dopamine release as a perceived saliency signal. These findings explain seemingly inconsistent data in the field that has suggested that dopamine is both a reward encoder and plays a critical role in neural responses to aversive and anxiogenic stimuli.[15,24,36,37,39,40] Our results recapitulate both data from experiments that test reward and aversive learning as well as present several additional experiments that dissociate valence-based predictive learning from perceived saliency. Our results provide a more complete framework of the role of dopamine in behavioral control.

These data rule out RPE coding by dopamine release signatures in the NAc core. In addition, we rule out several other forms of prediction error coding. We show that dopamine release is evoked to the omission of expected footshocks (Figure 4B); this is inconsistent with valence-free prediction error models, which hypothesize a negative response whenever a predicted outcome is omitted. Furthermore, using an experiment in which valence-free prediction error expects no change in response to the addition of a novel cue,[11,57] we showed a robust increase in dopamine release (Figure 4L). This rules out both valence-free and valence-based prediction error coding, as there is no change to the strength/valence of the previous cue-footshock association. In addition, we showed that dopamine release patterns in the NAc core do not fit unsigned prediction error, which signals an error in a prediction unidirectionally. Specifically, we showed that when an expected shock was omitted, the dopamine response was smaller than when the shock was presented as predicted rather than larger to signal the deviation from prediction (Figure 4B). Thus, dopamine release patterns in the NAc diverge from both RPE and other prediction error accounts in many contexts.

These studies are not the first to test the role of dopaminergic cells in prediction-based learning.[9,61–63] Many previous studies have focused on optical approaches to manipulate VTA activity to test whether behaviors that can be predicted by RPE are influenced by dopamine manipulations. While the results from several studies have been directly attributed to RPE coding by dopamine, some of these findings can also be explained by dopamine as a perceived saliency signal, which is highly influenced by novelty. For example, previous experiments have shown that VTA activity patterns follow cue-reward associations, and the optogenetic activation of VTA dopamine cell bodies elicits "unblocking," in which new learning is enabled for the additional cue.[9] This could be explained by increases in dopamine signaling deviations from prediction. However, these results are also expected if dopamine release is influenced by novelty, as novelty-induced unblocking has been reported.[64] Importantly, unblocking is not successful if novelty is reduced by pre-exposure to stimuli.[58,65] Therefore, even some previous evidence that has been deemed critical for the RPE account can be explained by dopamine responses to novelty.

Here, we provide an empirically verified computational model (Figure 7) that shows that mesolimbic dopamine release patterns conform to a perceived saliency signal. In this framework, the novelty and saliency of a stimulus are combined to dictate the allocation of attentional resources.[45,56,66,67] This model is not only capable of describing our own findings but also explains

Figure 7. Dopamine release signals perceived saliency

(A) The Kutlu-Calipari-Schmajuk (KCS) model. The model has 4 core components. (1) Associative component: based on a Rescorla-Wagner prediction error term. (2) Attentional component: mismatch between predicted/unpredicted stimuli increases novelty, and in turn, attention to all stimuli in the environment. (3) Perceived saliency: novelty, attention, and the physical intensity of a stimulus determine perceived saliency. (4) Behavioral response component: perceived saliency is combined with associative strength to produce an outcome.

(B–F) Experiments from Figure 4 were replotted to map onto model simulations.

(B) Switching from positive reinforcement to punishment (denoted by dotted line) decreased simulated and actual nose pokes ($r = 0.79$, $p < 0.0001$; $n = 5$).

(C) Perceived saliency of (gray) and dopamine responses to (blue) the cue similarly increased.

(D–F) Dopamine response to the cue increased after a shock was introduced ($t_4 = 2.76$, $p = 0.025$). KCS model simulations show that perceived saliency (E, $R = 0.67$, $p < 0.0001$) matches dopamine response patterns but prediction error does not (F, $r = 0.092$, $p = 0.23$).

(G) Dopamine release was stimulated from terminals in the NAc core via optogenetics.

(H) Dopamine release was evoked during fear conditioning to the cue on 25% of cue-shock pairings.

(I) Increasing NAc core dopamine decreased freezing compared to EYFP controls ($F_{1,20} = 17.84$, $p = 0.0004$; ChR2-Tone + Stim versus EYFP-Tone + Stim, $p \leq 0.0001$; ChR2-Tone + Stim versus ChR2-Tone only, $p \leq 0.0001$; $n = 5–6$) and non-stimulated trials in the same animals (ChR2-Tone + Stim versus ChR2-Tone only, $p = 0.0002$).

(J) KCS model simulations show that this behavior is predicted by increased perceived saliency.

(K) NAc core dopamine release was evoked at the time of the omitted shock during extinction.

(L) Dopamine stimulation prevented fear extinction in the ChR2 group compared to EYFP controls ($F_{1,9} = 5.90$, $p = 0.038$; last 4 trial block, $t_9 = 3.32$, $p = 0.0089$; $n = 5–6$).

(M) KCS simulations show that enhancing perceived saliency of the omitted shocks prevents extinction.

See Figures S6 and S7. Error bars represent ± S.E.M. * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$.

why dopamine can be mistaken for an RPE signal when results from specific experimental conditions are considered in isolation. Importantly, we directly replicate and explain data that have formed the basis of support for RPE; thus, our results are not in conflict with prior data, but do show that this widely accepted theory is fundamentally incomplete. This work extends our current understanding and provides a new formalized theory for the role of dopamine in learning and memory.

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
  - Apparatus
  - Surgical procedure
  - Histology
  - Fiber photometry
  - Fiber photometry analysis
  - Behavioral experiments
  - Positive reinforcement
  - Negative reinforcement
  - Positive punishment
  - Punishing behavioral responding by withholding/delaying the delivery of sucrose
  - Varying footshock intensities
  - Repeated neutral cue and footshock presentations
  - Varying sucrose and quinine delivery
  - Fear conditioning, omission, fear extinction
  - Introduction of a novel cue during fear conditioning
  - Optogenetic stimulation and inhibition of dopamine terminals
  - Machine learning
  - Video analysis and pose estimation via DeepLabCut
  - Computational modeling
  - Stimulus trace and value
  - Novelty
  - Attention
  - Aggregate stimulus prediction
  - Associative strength
  - Perceived saliency
  - Pavlovian conditioned response
  - Operant outcome
  - Operant response
  - Model simulations
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2021.08.052.

ACKNOWLEDGMENTS

This work was supported by NIH grants DA048931 and DA042111 to E.S.C., KL2TR002245 to M.G.K., GM07628 to J.E.Z., and DA045103 to C.A.S., as well as by funds from the VUMC Faculty Research Scholar Award to M.G.K.; the Pfeil Foundation to M.G.K.; Brain and Behavior Research Foundation to M.G.K., E.S.C., and C.A.S.; the Whitehall Foundation to E.S.C.; and the Edward Mallinckrodt, Jr. Foundation to E.S.C. The original computational model of Pavlovian conditioning that served as a framework for the KCS model was developed by Nestor A. Schmajuk, who passed away in 2015.

AUTHOR CONTRIBUTIONS

M.G.K. and E.S.C. conceptualized the study. M.G.K., J.E.Z., P.R.M., and S.J.K. performed surgeries and ran the behavioral and optogenetics experiments. M.G.K. and E.S.C. analyzed the data. M.G.K. and B.K. wrote the MATLAB codes for the analysis of machine learning, neural network modeling, and fiber photometry data. M.F.C. conducted the pose estimation (DeepLabCut) analysis. L.T. provided the dopamine sensor, dLight1.1. M.G.K., E.S.C., C.A.S., J.E.Z., P.R.M., and S.A.C. wrote the manuscript. All of the authors edited and approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 21, 2020
Revised: July 15, 2021
Accepted: August 18, 2021
Published: September 15, 2021

REFERENCES

1. Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.
2. Glimcher, P.W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. Proc. Natl. Acad. Sci. USA *108* (*Suppl 3*), 15647–15654.
3. Bayer, H.M., and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron *47*, 129–141.
4. Flagel, S.B., Clark, J.J., Robinson, T.E., Mayo, L., Czuj, A., Willuhn, I., Akers, C.A., Clinton, S.M., Phillips, P.E.M., and Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. Nature *469*, 53–57.
5. Iordanova, M.D. (2009). Dopaminergic modulation of appetitive and aversive predictive learning. Rev. Neurosci. *20*, 383–404.
6. Fiorillo, C.D., Newsome, W.T., and Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. Nat. Neurosci. *11*, 966–973.
7. Dayan, P., and Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. Curr. Opin. Neurobiol. *18*, 185–196.
8. Dayan, P., and Balleine, B.W. (2002). Reward, motivation, and reinforcement learning. Neuron *36*, 285–298.
9. 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.
10. Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. Nat. Rev. Neurosci. *17*, 183–195.
11. Rescorla, R., and Wagner, A. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In Classical Conditioning: Current Research and Theory, Vol. 2, A.H. Black, and W.F. Prokasy, eds. (Appleton-Century-Crofts).
12. Eshel, N., Tian, J., Bukwich, M., and Uchida, N. (2016). Dopamine neurons share common response function for reward prediction error. Nat. Neurosci. *19*, 479–486.

13. Takahashi, Y.K., Langdon, A.J., Niv, Y., and Schoenbaum, G. (2016). Temporal Specificity of Reward Prediction Errors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. Neuron *91*, 182–193.

14. Starkweather, C.K., Gershman, S.J., and Uchida, N. (2018). The Medial Prefrontal Cortex Shapes Dopamine Reward Prediction Errors under State Uncertainty. Neuron *98*, 616–629.e6.

15. 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.

16. Day, J.J., Roitman, M.F., Wightman, R.M., and Carelli, R.M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nat. Neurosci. *10*, 1020–1028.

17. Phillips, P.E.M., Stuber, G.D., Heien, M.L.A.V., Wightman, R.M., and Carelli, R.M. (2003). Subsecond dopamine release promotes cocaine seeking. Nature *422*, 614–618.

18. Howe, M.W., Tierney, P.L., Sandberg, S.G., Phillips, P.E.M., and Graybiel, A.M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature *500*, 575–579.

19. Saddoris, M.P., Sugam, J.A., Stuber, G.D., Witten, I.B., Deisseroth, K., and Carelli, R.M. (2015). Mesolimbic dopamine dynamically tracks, and is causally linked to, discrete aspects of value-based decision making. Biol. Psychiatry *77*, 903–911.

20. Saddoris, M.P., Cacciapaglia, F., Wightman, R.M., and Carelli, R.M. (2015). Differential Dopamine Release Dynamics in the Nucleus Accumbens Core and Shell Reveal Complementary Signals for Error Prediction and Incentive Motivation. J. Neurosci. *35*, 11572–11582.

21. Hart, A.S., Rutledge, R.B., Glimcher, P.W., and Phillips, P.E.M. (2014). Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. J. Neurosci. *34*, 698–704.

22. Bromberg-Martin, E.S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. Neuron *68*, 815–834.

23. Ungless, M.A., Magill, P.J., and Bolam, J.P. (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. Science *303*, 2040–2042.

24. Brischoux, F., Chakraborty, S., Brierley, D.I., and Ungless, M.A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. Proc. Natl. Acad. Sci. USA *106*, 4894–4899.

25. Valenti, O., Lodge, D.J., and Grace, A.A. (2011). Aversive stimuli alter ventral tegmental area dopamine neuron activity via a common action in the ventral hippocampus. J. Neurosci. *31*, 4280–4289.

26. McCutcheon, J.E., Ebner, S.R., Loriaux, A.L., and Roitman, M.F. (2012). Encoding of aversion by dopamine and the nucleus accumbens. Front. Neurosci. *6*, 137.

27. Redgrave, P., and Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? Nat. Rev. Neurosci. *7*, 967–975.

28. da Silva, J.A., Tecuapetla, F., Paixão, V., and Costa, R.M. (2018). Dopamine neuron activity before action initiation gates and invigorates future movements. Nature *554*, 244–248.

29. Howe, M.W., and Dombeck, D.A. (2016). Rapid signalling in distinct dopaminergic axons during locomotion and reward. Nature *535*, 505–510.

30. Parker, N.F., Cameron, C.M., Taliaferro, J.P., Lee, J., Choi, J.Y., Davidson, T.J., Daw, N.D., and Witten, I.B. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. Nat. Neurosci. *19*, 845–854.

31. Coddington, L.T., and Dudman, J.T. (2018). The timing of action determines reward prediction signals in identified midbrain dopamine neurons. Nat. Neurosci. *21*, 1563–1573.

32. 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.

33. Lewis, A.S., Calipari, E.S., and Siciliano, C.A. (2021). Toward Standardized Guidelines for Investigating Neural Circuit Control of Behavior in Animal Research. eNeuro *8*, ENEURO.0498-20.2021.

34. Kutlu, M.G., Zachry, J.E., Brady, L.J., Melugin, P.R., Kelly, S.J., Sanders, C., Tat, J., Johnson, A.R., Thibeault, K., Lopez, A.J., et al. (2020). A novel multidimensional reinforcement task in mice elucidates sex-specific behavioral strategies. Neuropsychopharmacology *45*, 1463–1472.

35. Patriarchi, T., Cho, J.R., Merten, K., Howe, M.W., Marley, A., Xiong, W.-H., Folk, R.W., Broussard, G.J., Liang, R., Jang, M.J., et al. (2018). Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. Science *360*, eaat4422.

36. Oleson, E.B., Gentry, R.N., Chioma, V.C., and Cheer, J.F. (2012). Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. J. Neurosci. *32*, 14804–14808.

37. Wenzel, J.M., Oleson, E.B., Gove, W.N., Cole, A.B., Gyawali, U., Dantrassy, H.M., Bluett, R.J., Dryanovski, D.I., Stuber, G.D., Deisseroth, K., et al. (2018). Phasic Dopamine Signals in the Nucleus Accumbens that Cause Active Avoidance Require Endocannabinoid Mobilization in the Midbrain. Curr. Biol. *28*, 1392–1404.e5.

38. Christianson, J.P., Fernando, A.B.P., Kazama, A.M., Jovanovic, T., Ostroff, L.E., and Sangha, S. (2012). Inhibition of fear by learned safety signals: a mini-symposium review. J. Neurosci. *32*, 14118–14124.

39. Menegas, W., Akiti, K., Amo, R., Uchida, N., and Watabe-Uchida, M. (2018). Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. Nat. Neurosci. *21*, 1421–1430.

40. de Jong, J.W., Afjei, S.A., Pollak Dorocic, I., Peck, J.R., Liu, C., Kim, C.K., Tian, L., Deisseroth, K., and Lammel, S. (2019). A Neural Circuit Mechanism for Encoding Aversive Stimuli in the Mesolimbic Dopamine System. Neuron *101*, 133–151.e7.

41. Stelly, C.E., Haug, G.C., Fonzi, K.M., Garcia, M.A., Tritley, S.C., Magnon, A.P., Ramos, M.A.P., and Wanat, M.J. (2019). Pattern of dopamine signaling during aversive events predicts active avoidance learning. Proc. Natl. Acad. Sci. USA *116*, 13641–13650.

42. Comoli, E., Coizet, V., Boyes, J., Bolam, J.P., Canteras, N.S., Quirk, R.H., Overton, P.G., and Redgrave, P. (2003). A direct projection from superior colliculus to substantia nigra for detecting salient visual events. Nat. Neurosci. *6*, 974–980.

43. Redgrave, P., Gurney, K., and Reynolds, J. (2008). What is reinforced by phasic dopamine signals? Brain Res. Brain Res. Rev. *58*, 322–339.

44. Pavlov, P.I. (2010). Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. Ann. Neurosci. *17*, 136–141.

45. Schmajuk, N.A., Gray, J.A., and Lam, Y.W. (1996). Latent inhibition: a neural network approach. J. Exp. Psychol. Anim. Behav. Process. *22*, 321–349.

46. Sokolov, E.N. (1990). The orienting response, and future directions of its development. Pavlov. J. Biol. Sci. *25*, 142–150.

47. Sokolov, E.N. (1978). A Neuronal Model of the Stimulus in a Reflex Arc. Soc. Psychol. *17*, 3–22.

48. Lubow, R.E. (1973). Latent inhibition. Psychol. Bull. *79*, 398–407.

49. Lubow, R.E. (1973). Latent inhibition as a means of behavior prophylaxis. Psychol. Rep. *32*, 1247–1252.

50. Lubow, R.E., Schnur, P., and Rifkin, B. (1976). Latent inhibition and conditioned attention theory. J. Exp. Psychol. Anim. Behav. Process. *2*, 163–174.

51. Lubow, R.E., and Moore, A.U. (1959). Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. J. Comp. Physiol. Psychol. *52*, 415–419.

52. Quintero, E., Díaz, E., Vargas, J.P., Schmajuk, N., López, J.C., and De la Casa, L.G. (2011). Effects of context novelty vs. familiarity on latent inhibition with a conditioned taste aversion procedure. Behav. Processes *86*, 242–249.

53. Sutton, R.S., and Barto, A.G. (1990). Time-derivative models of Pavlovian reinforcement. In Learning and Computational Neuroscience:

Foundations of Adaptive Networks., M. Gabriel, and J. Moore, eds. (The MIT Press), pp. 497–537.

54. Young, A.M.J., Joseph, M.H., and Gray, J.A. (1993). Latent inhibition of conditioned dopamine release in rat nucleus accumbens. Neuroscience *54*, 5–9.

55. 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.

56. Kutlu, M.G., and Schmajuk, N.A. (2012). Solving Pavlov's puzzle: attentional, associative, and flexible configural mechanisms in classical conditioning. Learn. Behav. *40*, 269–291.

57. Kamin, L.J. (1969). Predictability, surprise, attention and conditioning. In Punishment and Aversive Behavior, B.A. Campbell, and R.M. Church, eds. (Appleton-Century-Crofts).

58. Holland, P.C. (1999). Overshadowing and blocking as acquisition deficits: no recovery after extinction of overshadowing or blocking cues. Q. J. Exp. Psychol. B *52*, 307–333.

59. Brogden, W.J. (1947). Sensory pre-conditioning of human subjects. J. Exp. Psychol. *37*, 527–539.

60. Parker, K.E., Pedersen, C.E., Gomez, A.M., Spangler, S.M., Walicki, M.C., Feng, S.Y., Stewart, S.L., Otis, J.M., Al-Hasani, R., McCall, J.G., et al. (2019). A Paranigral VTA Nociceptin Circuit that Constrains Motivation for Reward. Cell *178*, 653–671.e19.

61. 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.

62. Kim, K.M., Baratta, M.V., Yang, A., Lee, D., Boyden, E.S., and Fiorillo, C.D. (2012). Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement. PLoS ONE *7*, e33612.

63. Chang, C.Y., Esber, G.R., Marrero-Garcia, Y., Yau, H.-J., Bonci, A., and Schoenbaum, G. (2016). Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. Nat. Neurosci. *19*, 111–116.

64. Dickinson, A., Hall, G., and Mackintosh, N.J. (1976). Surprise and the attenuation of blocking. J. Exp. Psychol. Anim. Behav. Process. *2*, 313–322.

65. Holland, P.C. (1985). Pretraining a compound conditioned stimulus reduces unblocking. Bull. Psychon. Soc. *23*, 237–240.

66. Mackintosh, N.J. (1975). A theory of attention: variations in the associability of stimuli with reinforcement. Psychol. Rev. *82*, 276.

67. 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.

68. Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat. Neurosci. *21*, 1281–1289.

69. Blaisdell, A.P., Bristol, A.S., Gunther, L.M., and Miller, R.R. (1998). Overshadowing and latent inhibition counteract each other: support for the comparator hypothesis. J. Exp. Psychol. Anim. Behav. Process. *24*, 335–351.

70. Grossberg, S. (1975). A Neural Model of Attention, Reinforcement and Discrimination Learning. In International Review of Neurobiology, C.C. Pfeiffer, and J.R. Smythies, eds. (Academic Press), pp. 263–327.

71. Grossberg, S., and Schmajuk, N.A. (1989). Neural dynamics of adaptive timing and temporal discrimination during associative learning. Neural Netw. *2*, 79–102.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-TH antibody	Millipore	#MAB318; RRID:AB_2313764
anti-GFP	Abcam	#AB13970; RRID:AB_300798
anti-chicken AlexaFluor 488	Life Technologies	#A-11039; RRID:AB_142924
anti-mouse AlexaFluor 594	Life Technologies	#A-21203; RRID:AB_141633
Bacterial and virus strains		
AAV_5.CAG.dLight1.1	UC Davis/Addgene	#111067-AAV5
AAV_5.hSyn.ChrimsonR-tdTomato	UNC vector core	N/A
AAV_9.rTH.PI.Cre.SV40	Addgene	#107788-AAV
AAV_5.Ef1a.DIO.eYFP	Addgene	#27056-AAV5
Experimental models: organisms/strains		
C57BL/6J	Jackson Laboratories	SN: 000664
Software and algorithms		
MATLAB R2019b	Mathworks	https://www.mathworks.com/products/new_products/release2019b.html
Prism9.0	Graphpad	https://www.graphpad.com/support/faq/prism-900-release-notes/
DeepLabCut	The Mathis Lab	http://www.mackenziemathislab.org/deeplabcut
Kutlu-Calipari-Schmajuk neural network model	This paper	https://github.com/kutlugunes/KCS_model.git
Support vector machine algorithms	This paper	https://github.com/kutlugunes/machine_learning

RESOURCE AVAILABILITY

Lead contact

Further information and requests for reagents and resources should be directed to and will be fulfilled by the Lead Contact, Dr. Erin S. Calipari (erin.calipari@vanderbilt.edu).

Materials availability

This study did not generate new unique reagents, plasmids, or mouse lines.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request
- All original codes have been deposited at Github and are publicly available as of the date of publication: https://github.com/kutlugunes.
- 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

Male and female 6- to 8-week-old C57BL/6J mice were obtained from Jackson Laboratories (Bar Harbor, ME; SN: 000664) and housed five animals per cage. All animals were maintained on a 12h reverse light/dark cycle. Animals were food restricted to 90% of free-feeding weight for the duration of the studies. Mice were weighed every other day to ensure that weight was maintained. All experiments were conducted in accordance with the guidelines of the Institutional Animal Care and Use Committee (IACUC) at Vanderbilt University School of Medicine, which approved and supervised all animal protocols. Experimenters were blind to experimental groups during behavioral experiments.

METHOD DETAILS

Apparatus

Mice were trained and tested daily in individual Med Associates (St. Albans, Vermont) operant conditioning chambers fitted with two illuminated nose pokes on either side of an illuminated sucrose delivery port, all of which featured an infrared beam break to assess head entries and nose pokes. One nose poke functioned as the active and the other as the inactive nose poke depending on the phase of the experiment (described below). Responses on the inactive nose poke were recorded but had no programmed consequence. Responses on both nose pokes were recorded throughout the duration of the experiments. Chambers were fitted with additional visual stimuli including a standard house light and two yellow LEDs located above each nose poke. Auditory stimuli included a white noise generator (used at 85 dB in these experiments) and a 16-channel tone generator capable of outputting frequencies between the range of 1 and 20 kHz (also presented at 85 dB).

Surgical procedure

Ketoprofen (5mg/kg; subcutaneous injection) was administered at least 30 mins before surgery. Under Isoflurane anesthesia, mice were positioned in a stereotaxic frame (Kopf Instruments) and the NAc core (bregma coordinates: anterior/posterior, +1.4 mm; medial/lateral, +1.5 mm; dorsal/ventral, −4.3 mm; 10° angle) or VTA (bregma coordinates: anterior/posterior, −3.16 mm; medial/lateral, + 0.5 mm; dorsal/ventral, −4.8 mm) were targeted (unilateral for fiber photometry and optogenetic stimulation experiments and bilaterally for optogenetic inhibition experiments). Ophthalmic ointment was applied to the eyes. Using aseptic technique, a midline incision was made down the scalp and a craniotomy was made using a dental drill. A 10-mL Nanofil Hamilton syringe (WPI) with a 34-gauge beveled metal needle was used to infuse viral constructs. Virus was infused at a rate of 50 nL/min for a total of 500 nL. Following infusion, the needle was kept at the injection site for seven minutes and then slowly withdrawn. Permanent implantable 2.5 mm fiber optic ferrules (Doric) were implanted in the NAc. Ferrules were positioned above the viral injection site (bregma coordinates: anterior/posterior, + 1.4 mm; medial/lateral, + 1.5 mm; dorsal/ventral, −4.2 mm; 10° angle) and were cemented to the skull using C&B Metabond adhesive cement system. Follow up care was performed according to IACUC/OAWA and division of animal care standard protocol. Animals were allowed to recover for a minimum of six weeks to ensure efficient viral expression before commencing experiments.

Histology

Subjects were deeply anesthetized with an intraperitoneal injection of Ketamine/Xylazine (100mg/kg;10mg/kg) and transcardially perfused with 10 mL of PBS solution followed by 10 mL of cold 4% PFA in 1x PBS. Animals were quickly decapitated, the brain was extracted and placed in 4% PFA solution and stored at 4 °C for at least 48-hours. Brains were then transferred to a 30% sucrose solution in 1x PBS and allowed to sit until brains sank to the bottom of the conical tube at 4 °C. After sinking, brains were sectioned at 35 μm on a freezing sliding microtome (Leica SM2010R). Sections were stored in a cryoprotectant solution (7.5% sucrose + 15% ethylene glycol in 0.1 M PB) at −20 °C until immunohistochemical processing. We immunohistochemically stained all NAc slices with an anti-GFP antibody (chicken anti-GFP; Abcam #AB13970, 1:200) for dLight1.1 and channelrhodopsin for the validation of viral placement. For the channelrhodopsin experiments, we also validated the targeting of TH+ cells in the VTA via an anti-TH antibody (mouse anti-TH; Millipore#MAB318, 1:100). Sections were then incubated with secondary antibodies [gfp: goat anti-chicken AlexaFluor 488 (Life Technologies #A-11039), 1:1000 and TH: donkey anti-mouse AlexaFluor 594 (Life Technologies # A-21203), 1:1000] for 2 h at room temperature. After washing, sections were incubated for 5 min with DAPI (NucBlue, Invitrogen) to achieve counterstaining of nuclei before mounting in Prolong Gold (Invitrogen). Following staining, sections were mounted on glass microscope slides with Prolong Gold antifade reagent. Fluorescent images were taken using a Keyence BZ-X700 inverted fluorescence microscope (Keyence), under dry 10x objective (Nikon). The injection site location and the fiber implant placements were determined via serial imaging in all animals.

Fiber photometry

For all fiber photometry experiments we injected the dopamine sensor dLight1.1 (AAV5.CAG.dLight1.1 (UC Davis)) into the NAc core. The fiber photometry recording system uses two light-emitting diodes (LED, Thorlabs) controlled by an LED driver (Thorlabs) at 490nm (run through a 470nM filter to produce 470nM excitation - the excitation peak of dLight1.1) and 405nm (an isosbestic control channel[35]). Light was passed through a number of filters and reflected off of a series of dichroic mirrors (Fluorescence MiniCube, Doric) coupled to a 400 μm 0.48 NA optical fiber (Thorlabs, 2.5mm ferrule size, optimized for low autofluorescence) and a 400 μm (0.48 NA) permanently implanted optical fiber in each mouse. LEDs were controlled by a real-time signal processor (RZ5P; Tucker-Davis Technologies) and emission signals from each LED stimulation were determined via multiplexing. The fluorescent signals were collected via a photoreceiver (Newport Visible Femtowatt Photoreceiver Module, Doric). Synapse software (Tucker-Davis Technologies) was used to control the timing and intensity of the LEDs and to record the emitted fluorescent signals. The LED intensity was set to 125 μW for each LED and was measured daily to ensure that it was constant across trials and experiments. For each event of interest (e.g., discriminative cue - S^d, headentries, shock, safety cue), transistor-transistor logic (TTL) signals were used to timestamp onset times from Med-PC V software (Med Associates Inc.) and were detected via the RZ5P in the synapse software (see below).

Fiber photometry analysis

The analysis of the fiber photometry data was conducted using a custom MATLAB pipeline. Raw 470nm (F470 channel) and isosbestic 405nm (F405 channel) traces were collected at a rate of 1000 samples per second (1kHz) and used to compute Δf/f values via polynomial curve fitting. For analysis, data was cropped around behavioral events using TTL pulses and for each experiment 2 s of pre-TTL and 18 s of post-TTL Δf/f values were analyzed. Δf/f was calculated as F470nm-F405nm/F405nm. This transformation uses the isosbestic F405nm channel, which is not responsive to fluctuations in calcium, to control for calcium-independent fluctuations in the signal and to control for photobleaching. Z-scores were calculated by taking the pre-TTL Δf/f values as baseline (z-score = (TTLsignal - b_mean)/b_stdev, where TTL signal is the Δf/f value for each post-TTL time point, b_mean is the baseline mean, and b_stdev is the baseline standard deviation). This allowed for the determination of dopamine events that occurred at the precise moment of each significant behavioral event. We also provided baselines as well as raw 405nm and 470nm traces used to calculate Δf/fs for the critical experiments as supplementary figures (Figure S2). For statistical analysis, we also calculated area under the curve (AUC) values for each individual dopamine peak via trapezoidal numerical integration on each of the z-scores across a fixed timescale which varied based on experiment. The duration of the AUC data collection was determined by limiting the AUC analysis to the z-scores between 0 time point (TTL signal onset) and the time where the dopamine peak goes back to baseline. AUC values were then normalized to the duration of the averaged peak (AUC value/AUC collection time in seconds) to avoid bias caused by data collection time.

Behavioral experiments

A series of behavioral experiments were run throughout this study to link dopamine responses to behavioral responding in Pavlovian and reinforcement contexts. They are outlined in detail below:

Positive reinforcement

Mice were trained to nose poke on an active nose poke for delivery of sucrose in a trial-based fashion. Following a correct response, the sucrose delivery port was illuminated for 5 s and sucrose was delivered (1 s duration of delivery, 10% sucrose w/v, 10ul volume per delivery). To create a trial-based procedure, a discriminative stimulus ($S^{d,\ sucrose}$) was presented signaling that responses emitted during the presentation of $S^{d,\ sucrose}$ resulted in the delivery of sucrose. Responses made during any other time in the session were recorded, but not reinforced. The discriminative stimulus was an auditory tone that consisted of 85dB at 2.5 kHz or white noise in a counterbalanced fashion. During the initial training, $S^{d,\ sucrose}$ was presented throughout the entirety of each 1-hour session and animals could respond for sucrose without interruption. When animals reached $\geq$ 80 active responses in a single session, they were then moved to a discrete trial-based structure in subsequent 1-hour sessions, wherein $S^{d,\ sucrose}$ was presented for 30 s at the beginning of each trial with a variable 30 s inter trial interval (ITI). Each trial ended following a correct response and associated sucrose delivery or at the end of a 30 s period with no active response. At the end of the trial both the trial and $S^{d,\ sucrose}$ were terminated. Animals that exhibited active responses in $\geq$ 80% of trials during a session then proceeded to the final phase of training wherein the duration of $S^{d,\ sucrose}$ was reduced to 10 s. Upon reaching the 80% criterion during this phase (i.e., acquisition), post-training dopamine responses were recorded over a 30 min session.

Negative reinforcement

Mice were trained to nose poke on the opposite, non-sucrose-paired nose poke for shock avoidance. Our previous studies showed the order of the positive and negative reinforcement training did not change the behavioral performance.[34] A second auditory discriminative stimulus ($S^{d,shock}$) - either tone or white noise, counterbalanced across animals and trial types - was presented at the beginning of each trial following a variable ITI as described above. In each trial the discriminative stimulus was presented for 30 s after which a series of 20 footshocks (1mA, 0.5 s duration) was delivered with a 15 s inter-stimulus interval. Trials ended when animals responded on the correct nose poke or at the end of the shock period. The end of the shock period was denoted by the presentation of a house light cue - termed safety cue - that signaled the end of the trial and was illuminated for one second. During these trials, mice could respond during the initial 30 s $S^{d,shock}$ period to avoid shocks completely, respond any time during the shock period to terminate the remaining shocks, or not respond at all. If mice did not respond both the trial and $S^{d,shock}$ were terminated after all 20 shocks had been presented (330 s total). Acquisition during negative reinforcement training was defined as receiving fewer than 25% of total shocks in a single one-hour session.

Positive punishment

Mice were first trained as described above for positive reinforcement. Following meeting acquisition criteria on positive reinforcement training (responding correctly on 100% of S^{d}) the contingency was switched to positive punishment where animals had to learn to inhibit responding to avoid footshock presentation. In these one-hour sessions, the trial structure was the same as in the positive reinforcement sessions (described above); however, in these trials responses on the previously sucrose-paired nose poke during the S^{d} resulted in immediate footshock delivery (1mA, 0.5 s duration) and the termination of the trial. A total of 15 positive reinforcement and 19 punishment training sessions were given during this task.

Punishing behavioral responding by withholding/delaying the delivery of sucrose

For the punishment experiment, animals received the same training as described above for sucrose reinforcement. Briefly, following a correct nose poke response during a 30sec discriminative stimulus ($S^{d,\ sucrose}$, 85dB white noise), the sucrose delivery port was illuminated for 5 s and sucrose was delivered (1 s duration of delivery, 10% sucrose w/v, 10ul volume per delivery). However, subsequently the task was changed so that the same $S^{d,\ sucrose}$ signaled that if animals responded sucrose would be withheld. The duration of the $S^{d,\ sucrose}$ was identical. In the case where mice made an operant response, the mice received no sucrose reward and the ITI period started. If the mice withheld their response for the duration of the $S^{d,\ sucrose}$ (30 s), sucrose delivery port was illuminated for 5 s and sucrose was delivered (1 s duration of delivery, 10% sucrose w/v, 10ul volume per delivery).

Varying footshock intensities

A total of 12 footshocks were delivered in a non-contingent and inescapable fashion over a 12-minute period. Shocks were delivered at 0.3mA, 1mA, and 1.7mA intensities (4 presentations for each shock intensity). Shocks were delivered in a pseudo-random order with variable inter-stimulus intervals (mean ITI = 30 s). All shock intensities were presented within the same test session.

Repeated neutral cue and footshock presentations

A total of 16 white noise stimuli (10 s in duration) were presented in a non-contingent fashion with a variable inter-stimulus interval (15, 30, or 45 s). All white noise stimuli were set to the same intensity (85dB). For the initial repeated footshock experiments, mice received 12 repeated 1mA footshocks with a fixed ITI (15 s). In order to test the effect of the saliency change on dopamine response following habituation, we presented a low intensity footshock (0.3mA) for 50 times with a fixed interval (15 s) and increased the intensity of the footshock for the 51st footshock presentation to 1mA. The high intensity footshock was presented once.

Varying sucrose and quinine delivery

Mice first completed positive reinforcement training (described above). In training sessions 10% sucrose (sucrose volume = 10ul) was delivered over a one second interval for all trials. In the testing sessions, following an active response sucrose was delivered for three seconds (high volume condition, 10uL/sec, sucrose volume = 30ul) on 50% of the trials or for only one second (low volume condition, 10uL/sec, sucrose volume = 10ul) for the remaining 50% of trials in a random fashion. For the quinine experiments, the mice were first trained to nose poke for sucrose as described above. During the test session, active responses resulted in the delivery of the bitter tastant quinine (0.03 g/l) for five seconds (high volume condition; 50% of trials) or one second (low volume condition; 50% of the trials).

Fear conditioning, omission, fear extinction

Mice received a single footshock (1mA, 0.5 s duration) immediately following a 5 s auditory cue (5kHz tone; 85dB) for 11 pairings. After a single conditioning session, mice underwent a session wherein 20% of shocks were omitted randomly after cue presentation followed by two extinction sessions in which the cue was presented, but shocks were omitted entirely.

Introduction of a novel cue during fear conditioning

An experiment was run to determine how the introduction of a novel cue altered dopamine responses to a previously shock-paired cue.[57] Animals were first trained based on traditional Pavlovian fear conditioning contingencies as described above. Mice received a single footshock (1mA, 0.5 s duration) immediately following a 5 s auditory cue (5kHz tone; 85dB) as described above. After 11 pairings, a novel cue (house light; 1 s duration) was presented concurrently at the onset of the auditory cue prior to footshock onset for 20% of trials at random.

Optogenetic stimulation and inhibition of dopamine terminals

Recording dopamine release while stimulating terminals via Chrimson

Mice were injected with AAV5.hSyn.ChrimsonR-tdTomato (UNC vector core) in the VTA. The same animals had dLight1.1 (AAV5.-CAG.dLight1.1) injected into the NAc core and a 400um fiber optic was implanted directly above the injection site. Because Chrimson is excited at 590nm and dLight at 470nm, this approach allows for simultaneous recording of dopamine release concomitant with stimulation of dopamine release from terminals within the same animals. For this series of experiments, mice received 10 footshocks (1mA, 0.5 s) before receiving the 11th footshock concurrent with the Chrimson stimulation 590nm, 1 s, 20Hz, 8mW). Stimulation was achieved via a yellow laser 590nm, which was modulated at 20Hz via a voltage pulse generator (Pulse Pal, Sanworks). The dopamine signal (via dLight) was recorded during the entire session. There were total of ten non-stimulated and one stimulated footshock presentations for each animal.

Optogenetic excitation of dopamine terminals via channelrhodopsin (ChR2)

In a separate group of C57BL/6J mice, AAV5.Ef1a.DIO.hchR2.eYFP (ChR2; UNC vector core) and AAV9.rTH.PI.Cre.SV40 (Addgene[60]) was injected into the VTA and a 200um fiber optic implant was placed into the NAc core. This allowed for the stimulation of dopamine release only in dopamine terminals that project from the VTA and synapse in the NAc core. Control animals received AAV5.Ef1a.DIO.eYFP injections into the VTA instead of ChR2. For these experiments, mice were trained with two sessions of cue-shock (1mA, 0.5 s) pairings (16 trials each). The cue was a 10 s tone (5 kHz at 85dB). On the third day, we delivered laser stimulation (473nm, 1 s, 20Hz, 8mW) into the NAc core at the onset of the tone for 25% of the trials (total 4 trials). The stimulation trials were

randomly intermixed within regular tone-shock trials (12 trials) with no laser stimulation. All trials ended with shock presentation.

For fear extinction experiments, following 3 sessions of fear conditioning, mice received an extinction session where 16 non-reinforced presentations of the tone cue in the absence of footshocks were given. All mice received blue laser stimulation (473nm, 1 s, 20Hz, 8mW) into the NAc core with the onset of the expected but now omitted footshock. We hand scored freezing behavior for the 10 s pre-footshock period for each trial in a blind fashion. The freezing response was defined as the time (seconds) that mice were immobile (lack of any movement including sniffing) during the tone period and calculated as percentage of total cue time.

Machine learning

The relationship between the dopamine signal obtained in our fiber photometry studies and the behavioral output during the positive/negative reinforcement experiments was analyzed using a support vector machine (SVM) classifier. A custom MATLAB code was used to create training and testing datasets for each dopamine signal associated with a behavioral outcome (e.g., response versus no response). We used the same number of pre-training, post-training as well as correct and missed trials between each comparison. Best predictive features for each unique signal-behavior tandem were extracted by employing sequential feature selection. The SVM model for best hyperparameters was trained using the best predictive features in a kernel function (radial basis function, RBF) to find the optimal hyperplane between binary prediction options (e.g., response versus no response). For the SVM classifier, we optimized the hyper parameters including C and gamma values using Bayesian optimization and an ''expected improvement'' acquisition function. Then the trained model was applied to the test dataset and prediction accuracy was calculated. All training and testing datasets were randomly selected. We repeated this process for 20 separate times for each dataset and reported accuracy as well as the number and type of correct and incorrect predictions in our dataset. Using the SVM classifier, we analyzed if the dopamine signal to pre- and post-training positive and negative reinforcement S^{d}s predicts response outcomes (response versus no response) within each trial. For negative reinforcement, we assumed all escape responses (responses made outside the initial 30 s shock free portion of the $S^{d,shock}$) as missed as all mice made an avoidance (response within the first 30 s of the $S^{d,shock}$) or escape response for each post-training trial. Then we analyzed whether the dopamine signal to the pre- versus post-training safety cue and shocks in the negative reinforcement paradigm predict future behavioral outcomes (response/avoidance versus no response/escape). For this analysis we only looked at the dopamine response to the first shock of each trial. We also tested whether the dopamine response to the last shock of each trial could predict the behavioral outcome on the next trial separately. We verified these results by using scrambled datasets where we obtained an average accuracy at chance (Figure S3).

Video analysis and pose estimation via DeepLabCut

We filmed the animals' movement using a USB camera (ELP, 1 megapixel) attached above the operant box. Videos were acquired at 10 frames/second and recorded through the integrated video capture within the Synapse software used for photometry. We used DeepLabCut (Python 3, DLC, version 2.2b8[68]) for markerless tracking of position. For fear conditioning experiments, DLC was trained on 4 top-view videos from 4 different animals. Twenty-five frames per animal (100 frames total) were annotated and used to train a ResNet-50 neural network for 200000 iterations. We used the location of the snout to compute the distance from the animal to the speaker and to compute movement. Trials in which the tracking quality was poor (< 0.9 likelihood score) during the test window were removed. The snout could not be tracked by DLC on a subset of frames because of a blind spot due the reflection of the infrared LED onto the plexiglass ceiling of the operant box. In these frames, we inferred the snout position as a smooth transition from the last position preceding the gap to the first following the gap.

For negative reinforcement experiments, DLC was trained separately for each animal using a side-view video. One hundred frames were annotated and used to train a ResNet-50 neural network for 200000 iterations. We used the location of the right ear to compute movement, as it was more easily traceable than the snout on side view videos. To quantify the orienting response to the light stimulus during negative reinforcement, the DLC network trained for fear conditioning was further refined by labeling 244 additional frames in which the change in brightness due to the light onset and offset deteriorated the accuracy of tracking. These frames were used to train the network used to analyze fear conditioning videos (top view) for an additional 80000 iterations. We computed the angle between the segment from the mouse's snout to the middle of its head (halfway between the ear) and the segment from the middle of the mouse's head to the light source. The angular speed, a proxy measure for the attention-like head movement at the time of the stimulus presentation, was calculated as the absolute change in angle over iterative frames.

Computational modeling

The Kutlu-Calipari-Schmajuk model (KCS; see Figure 7A for the model architecture) has been developed based on an attentional neural network model of Pavlovian conditioning (Schmajuk-Lam-Gray-Kutlu model, SLGK model[45,56]). At the core of the basic model (depicted in Figure 7A; see below for complete list of equations) is an error prediction term where associations between multiple conditioned stimuli (VCS-CS), as well as between conditioned and unconditioned (VCS-US) stimuli, are formed based on the same Rescorla-Wagner-based predictions. Thus, if dopamine does signal an RPE, it will still map onto this component within the model. However, the Rescorla-Wagner and other similar models rely entirely on the associative strength of the conditioned stimulus itself. As such, these models have two major weaknesses: 1. they do not account for the critical role of attention in associative learning and 2. they do not make predictions about the unconditioned stimulus itself and how a representation of this stimulus in both its presence and absence can contribute to behavioral control. To address these issues a term - ''Perceived Saliency''- is included to provide a representation of what is being predicted by a conditioned stimulus when the stimulus is physically present

as well as when it is absent. Perceived saliency focuses on the fact that the way an external stimulus is perceived by an organism is not only dependent on its physical properties (e.g., saliency), but rather is a combination of its physical and perceived saliency. The model assumes that perceived saliency of a stimulus is determined by stimulus saliency combined with the level of attention directed to a given stimulus. This way, two stimuli with equivalent physical intensity (e.g., two tones with equal dB values) can be weighted differently and receive processing priority when forming associations with an outcome depending on their attentional value.[69] Perceived Saliency is computationally defined as the product of stimulus saliency (termed 'CS' in the model) and attentional value of a stimulus (zCS). The core factor that controls attentional allocation is the level of Novelty in a given context, which is determined by the level of mismatch between predictions and actual occurrences of events on a global scale. Accordingly, the perceived saliency of a stimulus increases when Novelty is high and the organism directs more attention to that stimulus even when the saliency is constant. One of the most important tenets of the model is that even stimuli that are predicted but absent activate a representation, albeit weaker than the perceived saliency of stimuli that are present. The concept of novelty-driven perceived saliency allows our model to be able to describe learning phenomena where stimuli form associations with other stimuli in their absence [e.g., sensory preconditioning[59]; a type of learning which has been shown to be dependent on dopamine signaling[55]]

The constant values are determining rates of each term described below and are taken from the SLGK model (K1 = 0.2, K2 = 0.1, K3 = 0.005, K4 = 0.02, K5 = 0.005, K6 = 1, K7 = 2, K8 = 0.4, K9 = 0.995, K10 = 0.995, K11 = 0.75, K12 = 0.15, K13 = 4).

Stimulus trace and value

In the KCS model, time is represented as the units (t.u.) wherein stimuli are presented, making time-specific predictions of each component of the model possible. In addition to the duration of active presentation of a stimulus, the model assumes a short-term memory trace represented in time for each stimulus (τCS;[70,71]). The memory trace decays after the offset of the stimulus presentation. The strength of the memory trace of an individual stimulus (inCS) is determined by the conditioned stimulus saliency (CS) and strength of the prediction of the CS by other stimuli in the environment (preCS). K1 is the decay rate of the stimulus memory trace:

$$\tau CS = \tau CS + K1\ (CS - \tau CS)$$

$$inCS = \tau CS + K8\ preCS$$

Novelty

Novelty of a stimulus (S; CS or US) is proportional to the difference between the actual value (λS and isS) and the prediction of that stimulus (BS and ipreS) also denoted as the CS/US - preCS/US Integrator. Novelty increases when the stimuli are poor predictors of the US, other CSs (i.e., when the US, other CSs, or the context, CX, are underpredicted or overpredicted by the CSs and the CX). Total Novelty (Noveltytotal) is given by the Novelty values of stimuli present in the environment. The actual value of the US at a given time is proportional to the US in Pavlovian conditioning tasks and to Rout in operant conditioning tasks (see below for Operant outcome).

NoveltyS $\sim\Sigma$ |λS - BS|
Integrator = NoveltyS = |isS - ipreS|
isS = K9 isS+Rout (1 - isS)
ipreS = K10 ipreS + ipreS (1 - ipreS)
Noveltytotal = Noveltytotal + NoveltyS

Attention

Changes in attention zCS (ΔzCS,-1 > zCS > +1) to an active or predicted CS are proportional to the salience of the CS and are given by:

$$zCS = inCS\ ((K4\ OR\ (1 - |zCS|) - K5\ (1 + zCS)))$$

We assume that the orienting response (OR) is a sigmoid function of Novelty:

$$OR = \left(Noveltytotal^2 / \left(Noveltytotal^2 + K11^2\right)\right)$$

ΔzCS > 0; when Novelty > ThresholdCS
ΔzCS < 0; when Novelty < ThresholdCS

$$ThresholdCS = K5/K4$$

Aggregate stimulus prediction

The aggregate prediction of the US by all CSs with representations active at a given time (BUS) is determined by:

$$BUS = \Sigma XCSVCS - US$$

Associative strength

The change in the strength of an association between a CS and a US or between a CS and another CS is determined by:

$$\Delta VCS - US = K3\ XCS\ (\lambda US - BUS)\ |1 - VCS - US|$$

Perceived saliency

The Perceived Saliency of a CS or a US that is present or absent (given by the trace; τCS at a given time point), is proportional to the prediction of (BCS) and attention to that CS (zCS):

$$XCS \sim (\tau CS + BCS)\ zCS$$

$$XCS = K7\ (inCS\ K2 + (inCS\ zCS)$$

Pavlovian conditioned response

The US-specific CR is a sigmoid function of the total prediction of the US by all stimuli in the environment (preUS = preUS + BUS):

$$CR = K6\left(preUS^2 \Big/ \left(preUS^2 + K12^2\right)\right)(1 - OR)$$

Operant outcome

We assume that while the Pavlovian US takes a single value for each trial, the value of the operant outcome (Rout) is determined by the difference between two possible US outcomes (ΔUS), the US that occurs in the absence of the operant response (USno act) versus the US occurs following an operant response (USact):

$$\Delta US = (USact - USno\ act)\ K13$$

For positive reinforcement there are two potential USs:

The US in the absence of operant response = 0
The US following an operant response = +1

Then, ΔUS = USact - USno act = 1 - 0 = +1 for each operant response
For negative reinforcement there are two potential USs:

The US in the absence of operant response = -1
The US following an operant response = 0

Then, ΔUS = USact - USno act = 0 - (-1) = +1 for each operant response

Operant response

Our model assumes that the operant response is determined by the strength of the prediction of the operant outcome (Rout) in a probabilistic fashion:

if preUS $\leq$ 0.01 then (80%) Rout = USno act (no response) versus (20%) Rout = ΔUS (response)
if preUS > 0.01 then (20%) Rout = USno act (no response) versus (80%) Rout = ΔUS (response)

This way the preUS and Rout = ΔUS probability increases when ΔUS > 0 and decreases when ΔUS < 0.
For positive and negative reinforcement:

ΔUS = +1 therefore each response increases preUS and the Rout = ΔUS probability for next trial.

For positive and negative punishment:

ΔUS = -1 therefore each response decreases preUS and the Rout = ΔUS probability for next trial.

The Rout value was also used to calculate the prediction error (λUS = Rout), which was used to determine the VCS-US term described above:

$$Error = Rout - preUS$$

Model simulations

For each KCS model simulation, we determined 6 free parameters (ITI duration, cue duration, outcome duration, cue value, outcome value, and number of training trials) to mimic the experimental design of the behavioral experiments. Although these values are chosen arbitrarily, we kept them constant throughout the study. See below for the free parameters chosen for each experiment:

Operant ITI duration = 500 t.u.
Operant sucrose outcome value = 0.1
Operant shock outcome value = −0.1
Operant cue value = 1
Positive reinforcement cue duration = 40 t.u.
Positive reinforcement US duration = 40 t.u.
Negative reinforcement cue duration = 60 t.u.
Negative reinforcement safety cue duration = 5 t.u.
Negative reinforcement US duration = 5 t.u.
Pavlovian ITI duration = 200 t.u.
Pavlovian cue duration = 10 t.u
Pavlovian shock value = 1
Shock saliency shock values = 0.3, 1, 1.7
Pavlovian shock duration = 5 t.u.

For each operant conditioning simulation, we ran the simulations for 3 iterations and averaged the results. In the cases where the model failed to learn the task (< 5% of total simulations), we excluded the simulation values. For negative and positive reinforcement simulations, we assumed that pre-training lasted for 10 simulated trials. We assumed that the simulation met the learning criterion if it made an operant response for 10 consecutive trials. We used the last of these 10 trials for the post-training values. For all Pavlovian conditioning simulations, we assumed that the US value was the same regardless of the operant response condition.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using GraphPad Prism (version 8; GraphPad Software, Inc, La Jolla, CA) and MATLAB (Mathworks, Natick, MA). Z-scores were calculated as explained above (see Fiber photometry analysis). Unpaired and paired t tests and nested ANOVAs were employed for analysis where fiber photometry AUCs had two and three levels, respectively. Repeated-measures ANOVAs were used for the behavioral data from reinforcement studies (positive and negative reinforcement). For all repeated-measures ANOVA analysis, we used the Geisser-Greenhouse correction for sphericity. Independent t tests were used for analyzing machine learning studies where the accuracy was compared to a hypothetical mean (50% accuracy). We also calculated maximum z-scores for event fiber photometry traces and analyzed to determine if these were significantly different from the critical z-score at the p = 0.05 level (1.645) using independent-t tests. Alpha was set to 0.05 for all statistical analysis. Outliers were determined using the Grubbs's test for outliers (alpha = 0.0001). The exclusion criterion was established *a priori*. We assumed normal distribution of sample means for all t and F statistics. To compute the correlation between dopamine signal and movement, deltaF/F was z-scored on a trial-by-trial basis (10 s event-centered trials) and the mean value during the 2.5 s following cue onset was plotted as a function of the mean velocity for each trial. We used the function 'pearsonr()' (Python 3, Scypi 1.5.2) to compute the strength and significance of the correlation. A similar analysis was performed to compute the correlation following shock and to compute the correlation between dopamine and distance from snout to speaker following cue onset. For the correlations between angular speed and dopamine, we removed the trials in which the mouse was moving before stimulus onset to isolate and detect movement specific to the stimulus presentation. All data represented as mean ± SEM. Within the figures: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ and ns = not significant.