



Published in final edited form as:

Curr Biol. 2021 November 08; 31(21): 4748–4761.e8. doi:10.1016/j.cub.2021.08.052.

Dopamine release in the nucleus accumbens core signals perceived saliency

Munir Gunes Kutlu¹, Jennifer E. Zachry¹, Patrick R. Melugin², Stephanie A. Cajigas¹, Maxime F. Chevee¹, Shannon J. Kelley¹, Banu Kutlu¹, Lin Tian³, Cody A. Siciliano^{1,4}, Erin S. Calipari^{1,2,4,5,6,*}

¹Department of Pharmacology, Vanderbilt University, Nashville, TN, 37232, USA

²Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN 37232, USA

³Department of Biochemistry and Molecular Medicine, University of California, Davis, Sacramento, CA, 95817, USA

⁴Vanderbilt Center for Addiction Research, Vanderbilt University, Nashville, TN, 37232, USA

⁵Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, 37232, USA

⁶Department of Psychiatry and Behavioral Sciences, Vanderbilt University, Nashville, TN, 37232, USA

Summary:

A large body of work has aimed to define the precise information encoded by dopaminergic projections innervating the nucleus accumbens. Prevailing models are based on reward prediction error (RPE) theory where 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 nucleus accumbens (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-

*Corresponding Author/Lead contact: erin.calipari@vanderbilt.edu.

AUTHOR CONTRIBUTIONS

MGK and ESC conceptualized the study. MGK, JEZ, PRM, and SH performed the viral surgeries and ran the behavioral and optogenetics experiments. MGK and ESC analyzed the data. MGK and BK wrote the Matlab codes for the analysis of machine learning, neural network modeling, and fiber photometry data. MC conducted the pose estimation (deeplabcut) analysis. LT provided the dopamine sensor, dLight1.1. MGK, ESC, CAS, JEZ, PRM, and SC wrote the manuscript. All authors edited and approved the final version of the manuscript.

DECLARATION OF INTEREST

The authors report no conflicts of interest.

ADDITIONAL INFORMATION

All data in the manuscript or the supplementary material are available from the corresponding author upon reasonable request. Correspondence and requests for materials should be addressed to Erin S. Calipari. The Matlab codes used in this manuscript are available for public access at github.com/kutlugunes.

independent perceived saliency encoding across conditions. Together, 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 memory¹⁻⁷. The prevailing theory is that dopamine neurons projecting to the ventral striatum are the biological substrate for reward prediction error (RPE) where they transmit information about rewards and their predictive cues and update this information when errors in predictions are encountered^{8,9}. The prediction error hypothesis of dopamine signaling originates from an influential Pavlovian conditioning model, the Rescorla-Wagner model¹⁰, which assumes that learning occurs when the outcome is not perfectly predicted. Biological evidence for dopamine neurons encoding an RPE signal was first demonstrated by Schultz and colleagues¹ showing increases in dopamine neuron firing rates when an unexpected reward is encountered, a learning-dependent shift in firing to predictive 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 bodies¹¹⁻¹⁵ and in dopamine release in the nucleus accumbens (NAc)^{6,16-20}. As a result, the RPE model of dopamine's role in learning has been dominant in the field for the last 20 years. Nevertheless, there is literature suggesting that dopamine's role in learning and memory deviates from RPE²¹⁻³¹. In addition, 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.

The development of genetically encoded fluorescent dopamine sensors allows for direct, optical assessment of dopamine transients *in vivo*. To this end, we used the genetically encoded dopamine sensor, dLight1.1³², to record *in vivo* dopamine dynamics in the NAc (Figure S1A-D). Combining this approach with operant and Pavlovian tasks for a variety of stimuli - with both positive and negative valence - allows us to define how information is signaled via dopamine release in the NAc. First, we replicate work showing that dopamine release in the NAc core fits the canonical RPE model during learning reinforced by stimuli with positive valence. However, in tasks where learning was driven by stimuli with negative valence, NAc core dopamine deviates from RPE encoding. 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, prediction³³) and resultant dopamine signals, we first utilized the recently developed Multidimensional Cue Outcome Action Task (MCOAT)³⁴. We trained mice in a positive reinforcement task where

an auditory cue predicted that an operant response would result in delivery of sucrose [positive reinforcement] (Figure 1A,B). Subsequently, mice were trained in an aversive learning task where a distinct auditory cue indicated that a different response would prevent the delivery of footshocks [negative reinforcement] (Figure 1I,J). The operant response in the two phases requires the same motoric action (nose poke), and correct responses both 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 the neural signatures should be similar when positive and negative reinforcement tasks are successfully performed.

During positive reinforcement (Figure 1A), we recorded NAc core dopamine responses to an auditory discriminative cue ($S^d, \text{sucrose}$) and the contingent delivery of sucrose. This was done at two timepoints: during the first training session and in the same mice after they had met acquisition criteria (post-training; >80% correct responses) (Figure 1B,C; also see Figure S1E,F). Dopamine responses to the $S^d, \text{sucrose}$ increased from pre- to post-training whereas dopamine responses during the positive outcome (sucrose retrieval) decreased (Figure 1D-H; see Figure S1K for additional analyses). In agreement with previous literature^{1,12,15}, this pattern follows what is suggested by the Rescorla-Wagner model and RPE based accounts of dopamine in reward-based predictions (Figure S1G,H).

During negative reinforcement, a separate auditory cue (S^d, shock) was presented which signaled the ability to nose poke to avoid a series of footshocks (Figure 1I,J). Correct responses resulted in avoidance or termination of shocks in each trial. A cue light (termed ‘safety cue’) signaled the end of the shock period regardless of whether this occurred through an operant response or time elapsed (Figure 1I).

Mice showed robust negative reinforcement learning (Figure 1K; also see Figure S1I,J). Unlike positive reinforcement, dopamine responses to the S^d, shock for negative reinforcement did not increase with training (Figure 1L,M,N,Q). The dopamine response to the safety cue (a positive outcome) was largest pre-training and was reduced with experience (Figure 1O,Q; see Figure S1L for additional analyses). Importantly, footshocks evoked a positive dopamine response (Figure 1L,M,P,Q), which was not due to movement (Figure S2A-S2G). Moreover, the dopamine response to aversive footshocks increased as animals’ performance in the task increased (Figure 1P,Q).

These data are particularly surprising as earlier valence-based prediction error accounts (e.g., RPE) considered dopamine responses to safety cues as a part of reward processing³⁵. However, previous studies primarily analyzed the dopamine response to safety cues after it gained positive valence, rather than over the entire learning process. It is important to note that while safety cues ultimately serve as a reward, the cue itself is not inherently rewarding and it only acquires value after the animal learns that it predicts the removal/avoidance of an aversive event³⁶. These data show that the largest dopamine response was when the safety cue is novel and before valence could have been attributed. Additionally, in contrast to the 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

the footshock. Lastly, unlike in positive reinforcement, dopamine responses to the outcome (footshock) increase with training instead of decreasing.

Dopamine responses to an aversive outcome, but not safety cues, predicts future behavior

In previous studies, the results showing that the dopamine signal to the safety cue, a positive outcome, goes down as the animals learn the task have been interpreted as being in line with RPE models^{35,37-40}. While the data presented above suggest that this is likely not the case, if the RPE interpretation is correct, the dopamine signal to the safety cue would serve as an error signal and should function to update subsequent decisions. To test this, we employed a supervised machine learning approach to examine if features of the dopamine signal to cues and outcomes on a trial-by-trial basis are predictive of behavioral performance. We used a Support Vector Machine (SVM) where we iteratively divided our data into training and test sets to determine if 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 the animal made a response or not in the current trial based on the dopamine response to the $S^{d, \text{sucrose}}$ presented during positive reinforcement with ~80% accuracy (Figure 2B-E). However, SVM was not able to predict behavioral responses based on the dopamine response to the $S^{d, \text{shock}}$ predicting negative reinforcement trials (Figure 2F-I; also see Figure S3 for additional analyses). 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 $S^{d, \text{shock}}$ on the next trial or not; Figure 2J and Figure S3G,H and also see Figure S3I-K for additional analyses), 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.

Finally, the SVM was unable to predict the response on the subsequent trial above chance (~50% accuracy; Figure 2K and Figure S3L,M) indicating that this signal was not predictive of future behavior. Importantly, after training, dopamine responses to safety cues were larger in trials where a change in future behavior was not necessary (avoidance trials) compared to those where updating would improve performance (escape trials) (Figure S3N-P). Therefore, contrary to the interpretation that the dopamine response to the safety cue serves as a "positive" outcome, we found that this signal does not track errors needed to update future behavior.

These data suggest that dopamine responses track novel and salient events, rather than valence-based predictions. Therefore, we designed a series of studies to systematically modulate levels of the physical intensity of stimuli and/or novelty in the environment to determine if dopamine scaled with these factors, or if dopamine scaled more accurately with prediction-based learning algorithms.

Accumbal dopamine tracks stimulus saliency

While the data above show that dopamine does not signal RPE, dopamine release could still drive behavior by encoding other necessary components of associative learning. The quantifiable physical intensity of a stimulus (e.g., amperage, decibel, lux), termed saliency, is a principal component of contingency learning^{1,10}. To probe how dopamine release

changes to scaling unconditioned stimulus intensity, we recorded dopamine responses to random shocks of varying intensities (Figure 3A).

Dopamine responses to the shock increased as intensity was increased (Figure 3B,C). 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 tracked sucrose volume (Figure 3D,E). Additionally, we exposed animals to the bitter tastant quinine and also observed larger dopamine responses to higher volumes of quinine (Figure 3F,G). Therefore, dopamine release tracked in the same direction (i.e., increase) for stimuli with both positive (sucrose) and negative (shock, quinine) valence. These results suggest that NAc core dopamine is influenced by stimulus intensity.

NAc core dopamine does not follow prediction error 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). To test these competing hypotheses, mice went through a series of fear conditioning sessions where the probability of shock presentation was altered to determine how dopamine responses to the conditioned cue and unconditioned stimulus changed over acquisition, during unexpected omissions, and over extinction.

Mice were trained in a single fear conditioning session, where they received 11 tone-footshock pairings (Figure S4A). During this initial session, tone was followed by shock for 100% of trials. We found that the aversive footshocks resulted in a robust positive dopamine response (Figure 4A,B; also see Figure S4A,B) and this signal stayed positive throughout the fear conditioning session (Figure S4A-C).

After initial training, shock was delivered on 80% of the trials and was omitted on the remaining 20% (Figure 4A-E). There were no differences in freezing to the fear cues (Figure 4A) or dopamine response to the tone (Figure S4D,E) during omission trials. We found that the omission of the footshock resulted in a positive response at the time of the expected footshock, though the omission response was smaller than when the shock was present (Figure 4B). 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. Interestingly, we also found that the dopamine response to the footshock was stronger during the omission session where the footshocks were presented on only 80% of the trials compared to the first fear conditioning session where shocks were presented with 100% probability, likely because the outcome is less certain and thus perceived as more salient when it does occur (Figure 4D,E).

Next, we examined the dopamine response to predictive cues and footshocks during extinction (Figure 4F-J). The freezing response to the fear cues progressively decreased with fear extinction (Figure 4F); however, the dopamine response to the predictive cues did not change (Figure 4G,H), demonstrating – along with additional studies with appetitive cues (Figure S4F-K) – that cue responses are not a function of learned negative valence (also see Figure S4L-Q). Additionally, the NAc core dopamine signal to the footshocks was strongest during fear conditioning and decreased over extinction when animals learned that the conditioned stimulus no longer predicted the presentation of a footshock (Figure 4I,J). This pattern further ruled out both signed and unsigned prediction error accounts as described above.

Lastly, 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 (Figure 4K-O). Mice were trained to associate an auditory cue (tone) with the presentation of a footshock. Once this was learned, a novel cue (light) was presented concurrent with the original cue followed by a footshock. In this case, 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, it is a novel stimulus. Importantly, novel stimuli can decrease conditioned responses when presented together with a learned conditioned stimulus, a phenomenon known as external inhibition⁴¹. This is one of the main challenges to the Rescorla-Wagner learning model which cannot account for the change in conditioned response given that no error in the original prediction has occurred^{42,43}.

When a novel light stimulus was presented with the fear cue, freezing was reduced compared to the trials where the fear cue was presented alone (Figure 4K), demonstrating that the novel cue produced 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 (Figure 4L,M). Further, the dopamine response to the footshocks did not differ across trial types (Figure 4N,O). Thus, these data show that dopamine responds to valence-free stimuli that do not signal a change in associative strength. Together, these data show that dopamine consistently tracks stimulus features related to saliency across conditions.

NAc core dopamine responds to stimulus saliency and is dynamically modulated by novelty during learning.

Our results show that physical intensity of stimuli determines the dopamine response, but this explanation 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 saliency⁴⁴⁻⁴⁶. Perceived saliency contributes to learning and can change over experience, and therefore could provide a parsimonious explanation to seemingly disparate results showing that dopamine signals physical stimulus saliency but also is modulated across learning^{47,48}.

Perceived saliency is highly influenced by the novelty of the environment/stimulus. Novelty is determined by the level of pre-exposure to a stimulus/context, where novelty is high when

a stimulus is encountered for the first time and decreases with repeated exposure^{47,49,50}. If NAc core dopamine tracks perceived saliency it should initially track physical intensity, as demonstrated above (Figure 3), but also be dynamically influenced by novelty (as observed in Figure 4K-O); thus, we hypothesized that dopamine release would decrease with the repeated presentation of a stimulus despite the valence and intensity being held constant. Thus, we presented repeated footshocks with the same physical intensity (1.0 mA) at fixed time intervals while recording dopamine responses over a single test session (Figure 5A). Supporting our hypothesis, the amplitude of dopamine release to the footshocks decreased progressively with repeated presentations (Figure 5A,B). This effect was not due to detection limits of shock-evoked dopamine release (Figure S5A-E) or potential changes in baseline fluorescence over the recording session and was insensitive to the methodology for quantifying transient magnitude (Figure S5F-J). Further, this pattern was observed whether the footshocks were presented in a random fashion (Figure 5A,B) or in a negative reinforcement context (Figure S5K-Q). Therefore, as footshocks become less novel, thereby decreasing perceived saliency, the magnitude of the dopamine signal decreases.

We next sought to parse if perceived saliency encoding could be observed in scenarios that can be fully described by RPE. We conducted an experiment where mice nose poked on a fixed-ratio one schedule for the presentation of a reward (sucrose) following the presentation of an auditory discriminative stimulus (S^d ; Figure 5C). Following acquisition, the outcome was switched so that a nose poke in the presence of the S^d resulted in a footshock (and no sucrose delivery) while all other task parameters remained the same. 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 presentation and responses to the predictive cue should decrease as the value is updated. Alternatively, 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, the response to the predictive cue should also increase as the familiar cue acquires a novel prediction.

Following the contingency switch, mice decreased and eventually stopped responding (Figure 5D) demonstrating that the shock effectively functioned as a punishment 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 robust positive dopamine transient (Figure 5E). Further, after just a single trial, the dopamine response to the next S^d presentation following the unexpected shock was increased in magnitude even though the cue was familiar, presented at a fixed intensity, and now predicted a negative outcome (Figure 5F-H). These results demonstrate that dopamine release is increased when a novel stimulus is introduced (even when there is a negative error in reward prediction) 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. We analyzed the first presentations of tone, white noise, and house light in these experiments – before animals were able to make any behavioral responses or learn about these stimuli. These neutral stimuli evoke dopamine release (Figure 6A-C). Further, a neutral white noise stimulus in the absence of any outcome evoked a positive dopamine transient upon presentation, which decreased over repeated exposure – as would be predicted for a perceived saliency signal (Figure 6D-F).

Next, we repeatedly presented a mild footshock (0.3mA, 50 presentations). As we showed above, the dopamine response habituated with repeated presentations (Figure 6G,H) but the additional trials in this experiment revealed that the dopamine response to footshocks decreased to baseline, again demonstrating that physical intensity alone cannot explain dopamine responses. On the 51st presentation, we increased the intensity of the footshock to 1mA, which had not been experienced prior and is therefore both salient and novel. The dopamine response to shock was again observed, despite being habituated to the 0.3mA shock just prior (Figure 6I,J). Further, although the physical intensity was now stronger, the response was not larger than the response to the initial 0.3 mA shock at the beginning of the session (Figure 6K,L). This demonstrates that perceived saliency, which is a product of both novelty and physical intensity, can explain NAc dopamine release patterns across a wide range of conditions.

A novel model of behavioral control: the Kutlu-Calipari-Schmajuk 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 of the narrow parameters they were originally designed to explain. Even more recent accounts of Pavlovian conditioning, such as the temporal difference model^{26,51}, fall short as they cannot explain concepts like latent inhibition and sensory preconditioning that have been shown to be directly altered by dopamine manipulations^{52,53}. By integrating several critical theoretical constructs (i.e., prediction error, association formation, attention, and temporal dynamics), we developed a new behavioral model, the Kutlu-Calipari-Schmajuk (KCS) model, that allows us to explore the involvement of dopamine in both Pavlovian and operant learning. Importantly, the KCS model is a model of behavior, which allows for unbiased mapping of dopamine onto its theoretical components across contexts and conditions in a data-driven fashion.

We based our model on an attentional neural network model of Pavlovian conditioning^{46,54}. At the core of the model (depicted in Figure 7A; see **Materials and Methods** for complete list of equations) is an error prediction term where associations are formed based on Rescorla-Wagner-based predictions. An additional critical aspect of this model is *perceived saliency* which 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., intensity), but is highly influenced by the level of attention directed to that stimulus. Perceived saliency 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, the perceived saliency of a stimulus increases when novelty is high,

and the organism directs more attention to that stimulus even when the physical intensity is constant.

Together, this model of behavior can replicate Pavlovian conditioning paradigms⁵⁴ including basic learning phenomena that the Rescorla-Wagner model was developed to describe, such as blocking⁵⁵ and overshadowing⁵⁶; Figure S6A) as well as many that it cannot⁴⁷; Figure S6B,D-F). It can also describe temporal learning phenomena⁵¹; Figure S6C) and basic operant conditioning schedules (Figure S6D-F).

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 (Figure 4; Figure S6 and S7). The KCS model successfully predicted behavioral response patterns during fear conditioning, footshock omission, fear extinction, and introduction of a novel stimulus (Figure S7A-F). The perceived saliency component of the KCS model alone tracked dopamine patterns in all cases (Figure S6G-M and S7G-L).

Further, we simulated the experiment from Figure 5C-H where 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 (Figure 7C, D) which follow changes in perceived saliency but are opposite from prediction error simulations (Figure 7E,F). Together, these data show the dopamine maps onto perceived saliency across the conditions tested within this study.

Testing model predictions via optogenetic manipulation of dopamine release

Using the KCS model to guide experimental design we ran two optogenetic experiments to test whether dopamine transmits a perceived saliency signal. 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 controls⁵⁷] in the VTA (Figure 7G). By implanting a fiber optic over the NAc core, we selectively stimulated NAc core dopamine terminals (Figure S6N).

In Figure 4K-M, 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 error in the previously learned prediction. We hypothesized that if dopamine signals perceived saliency that increasing dopamine to a previously learned fear cue would induce external inhibition and decrease freezing during those trials. We conducted an experiment where following an initial fear conditioning training session, we stimulated dopamine terminals in the NAc core during 25% of the cue presentations (Figure 7H). Indeed, when dopamine was artificially increased during the cue, freezing was decreased compared to controls, including the cue alone trials in the same session (Figure 7I). When perceived saliency values for fear conditioning cues were computationally increased using the KCS model, the predicted conditioned response was similarly decreased (Figure 7J); thus, confirming that dopamine transmits a perceived saliency signal.

Above, we observed a dopamine response at the time of a predicted but absent footshock in Figure 4A-C. We hypothesized that if this dopamine signal signals perceived saliency that optogenetically enhancing it should prevent extinction (Figure 7K). Indeed, stimulating dopamine at the time of the omitted shock prevented extinction (Figure 7L). Similarly, computationally amplifying the perceived saliency of the omitted outcomes within the KCS model also similarly interferes with extinction learning (Figure 7M). Together, these data confirm that NAc core dopamine signals perceived saliency.

DISCUSSION

We show that dopamine responses track the perceived saliency of stimuli across all conditions, learning schedules, and contexts. 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 data are critically important as they explain seemingly inconsistent data in the field that has suggested that dopamine is both a reward encoder and plays a critical role in aversive learning and neural responses to anxiogenic stimuli^{1,12,29,35,37-39}. Together, 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. Combined, our results provide a more complete framework of dopamine's role in behavioral control.

The results presented here as well as results from others⁵⁸⁻⁶⁰ rule out reward-specific processing by dopamine in the NAc, as dopamine release is also elicited by aversive events and neutral stimuli. We also rule out other forms of prediction-error coding. First, we showed that dopamine release is evoked to the omission of expected footshocks, while valence-free prediction error would predict a negative dopamine response in this case. Further, using an experiment where valence-free prediction error expects no change in response to the addition of a novel light cue^{10,55}, we showed a robust increase in dopamine release to this novel neutral cue. This result rules out both the valence-free and valence-based prediction error encoding as there is no change to the strength or 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 the different than expected outcome unidirectionally. Specifically, we showed that when an expected shock was omitted the dopamine response was weaker than in the case where the expected shock was presented. Thus, dopamine release patterns in the NAc diverge from both RPE and prediction error accounts in many contexts.

These studies are not the first to test the role of dopaminergic cells in prediction-based learning. Many previous studies have focused on optical approaches to manipulate VTA cell body activity to test whether behaviors that can be predicted by RPE computations are influenced by dopamine manipulations. While the results from some previous studies have been directly attributed to RPE encoding 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 studies examining the role of dopamine in blocking experiments have shown that VTA activity patterns follow cue-reward associations

and optogenetic activation of VTA dopamine cell bodies elicits “unblocking” where new learning was enabled for the additional cue⁸. 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. Novelty-induced unblocking, which is achieved when novelty is increased through outcome omission, has been reported⁶¹. Further, unblocking, through the increase of stimulus magnitude is not successful if novelty is reduced by pre-exposure to the stimulus⁶², or previously learning that the stimulus does not predict an outcome⁶³. Therefore, even some previous evidence that has been deemed critical for the reward prediction error account can ultimately be explained by dopamine responses to novelty in the environment.

We show here that NAc dopamine signals perceived saliency, where the novelty and saliency of a stimulus are combined to dictate the allocation of attentional resources. For example, repeated presentations of footshocks remain salient but less novel for the animal resulting in a decreased dopamine response. Similarly, omission of footshocks is more novel but it is presumably a less salient experience than receiving a footshock that is expected - both events evoke a significant positive dopamine response in the NAc core. Further, when operant outcomes were changed from a sucrose reward to footshocks unexpectedly, the dopamine response to the cues predicting these outcomes also increased, which is a novel/salient experience, but worse than expected outcome. Theoretical constructs like perceived saliency as a product of attentional value have been conceptualized by previous associative learning models^{46,54,64,65}.

This interpretation is also consistent with work that has suggested that dopamine activity is critically linked to saliency encoding, especially as it relates to switching attention to unexpected and behaviorally important stimuli⁶⁶. It is important to note that this does not negate the role of dopamine in prediction-based learning, as the perceived saliency of a stimulus functions to allow for accurate predictions in changing environments. In fact, prediction error signaling, and the fundamental constructs of perceived saliency, novelty, and saliency, are intrinsically linked and often go in the same direction, explaining why dopamine release in reward-based contexts maps onto both. Only in situations where prediction error and perceived saliency are dissociable do diversions between the two computations emerge.

By experimentation that integrates learning across contexts, behavioral action, and valence, we provide a comprehensive and formalized framework for accumbal dopaminergic release in valence-independent behavioral control. We conclude that NAc core dopamine release tracks the perceived saliency of external stimuli and events instead of prediction error across a variety of contexts and conditions to drive adaptive behavior in all cases.

STAR METHODS

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.

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 (which were 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 .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 DAC 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 anaesthetized 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 35µm 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 Irvine)) 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 54). Light was passed through a number of filters and reflected off of a series of dichroic mirrors (Fluorescence MiniCube, Doric) coupled to a 400µm 0.48 NA optical fiber (Thorlabs, 2.5mm ferrule size, optimized for low autofluorescence) and a 400µm (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 125µW 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 - Sd, 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 f/f values via polynomial curve fitting. For analysis, data was cropped around behavioral events using TTL pulses and for each experiment 2s of pre-TTL and 18s of post-TTL f/f values were analyzed. f/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 f/f values as baseline ($z\text{-score} = (TTL\text{signal} - b_mean)/b_stdev$, where TTL signal is the f/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 f/f s 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.

METHOD DETAILS

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 – denoted by its illumination – for delivery of sucrose in a trial-based fashion. Following a correct response, the sucrose delivery port was illuminated for 5 seconds and sucrose was delivered (1s duration of delivery, 10% sucrose w/v, 10ul volume per delivery). To create a trial-based procedure, a discriminative stimulus (Sd, sucrose) was presented signaling that responses emitted during the presentation of Sd, 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, Sd, sucrose was presented throughout the entirety of each 1-hour session and animals could respond for sucrose without interruption. When animals reached 80 active responses in a single session, they were then moved to a discrete trial-based structure in subsequent 1-hour sessions,

wherein Sd, sucrose was presented for 30 seconds at the beginning of each trial with a variable 30 second inter trial interval (ITI). Each trial ended following a correct response and associated sucrose delivery or at the end of a 30 second period with no active response. At the end of the trial both the trial and Sd, sucrose were terminated. Animals that exhibited active responses in 80% of trials during a session then proceeded to the final phase of training wherein the duration of Sd, sucrose was reduced to 10 seconds. 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 did not change the behavioral performance 58. A second auditory discriminative stimulus (Sd, 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 seconds after which a series of 20 footshocks (1mA, 0.5 second duration) was delivered with a 15 second 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 was illuminated for one second. During these trials, mice could respond during the initial 30 second Sd, 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 Sd, shock were terminated after all 20 shocks had been presented (330 seconds 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 Sd, sucrose) 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 Sd, sucrose resulted in immediate footshock delivery (1mA, 0.5 second 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 a discriminative stimulus (Sd, sucrose, 85dB white noise), the sucrose delivery port was illuminated for 5 seconds and sucrose was delivered (1s duration of delivery, 10% sucrose w/v, 10ul volume per delivery). However, subsequently the task was changed so that the same Sd, sucrose signaled that if animals responded sucrose would be withheld. The duration of the Sd, 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 Sd, sucrose (30 seconds), sucrose delivery port was illuminated for 5 seconds and sucrose was delivered (1s 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 sec). All shock intensities were presented within the same test session.

Repeated neutral cue and footshock presentations: A total of 16 white noise stimuli (10 seconds in duration) were presented in a non-contingent fashion with a variable inter-stimulus interval (15, 30, or 45 seconds). All white noise stimuli were set to the same intensity (80dB). For the initial repeated footshock experiments, mice received 12 repeated 1mA footshocks with a fixed ITI (15 seconds). 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 seconds) 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 animals that were trained to nose poke for sucrose as described above were given the bitter tastant quinine (0.03g/l) as during the test session. Specifically, following an active response during the Sd, quinine, the quinine solution was delivered 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 second duration) immediately following a 5 second 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 76. Animals were first trained based on traditional Pavlovian fear conditioning contingencies as described above. Mice received a single footshock (1mA, 0.5 second duration) immediately following a 5 second auditory cue (5kHz tone; 85dB) as described above. After 11 pairings, a novel cue (house light; 1 second 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 sec) before receiving the 11th footshock concurrent with the Chrimson stimulation 590nm, 1s, 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 10 non-stimulated and 1 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; 78) 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 sec) pairings (16 trials each). The cue was a 10 sec tone (5 kHz at 85dB). On the 3rd day, we delivered laser stimulation (473nm, 1s, 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, 1s, 20Hz, 8mW) into the NAc core with the onset of the expected but now omitted footshock. We hand scored freezing behavior for the 10 second 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 data sets for each dopamine signal associated with a behavioral outcome (e.g., response vs. 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 vs. 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 data set 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 Sds predict response outcomes (response versus. no response) within each trial. For negative reinforcement, we assumed all escape responses (responses made outside the initial 30 second shock free portion of the Sd,shock) as missed as all mice made an avoidance (response within the first 30 seconds of the Sd,shock) or escape response for each post-training trial. Then we analyzed whether the dopamine signal to the pre- vs. 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 data sets 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, 112) 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; 57,68). At the core of the basic model (depicted in Figure 6A; 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 113. 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; 114, a type of learning which has been shown to be dependent on dopamine signaling; 75].

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 (τ_{CS} ; 115,116). The memory trace decays after the offset of the stimulus presentation. The strength of the memory trace of an individual stimulus (in_{CS}) is determined by the conditioned stimulus saliency (CS) and strength of the prediction of the CS by other stimuli in the environment (pre_{CS}). K_1 is the decay rate of the stimulus memory trace:

$$\tau_{CS} = \tau_{CS} + K_1 (CS - \tau_{CS})$$

$$in_{CS} = \tau_{CS} + K_8 pre_{CS}$$

Novelty: Novelty of a stimulus (S; CS or US) is proportional to the difference between the actual value (λ_S and is_S) and the prediction of that stimulus (BS and $ipre_S$) also denoted as the CS/US - pre_{CS}/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 ($Novelty_{total}$) 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 R_{out} in operant conditioning tasks (see below for Operant outcome).

$$Novelty_S \sim \Sigma \lambda_S - BS$$

$$Integrator = Novelty_S = is_S - ipre_S$$

$$is_S = K_9 is_S + R_{out} (1 - is_S)$$

$$ipre_S = K_{10} ipre_S + ipre_S (1 - ipre_S)$$

$$Novelty_{total} = Novelty_{total} + Novelty_S$$

Attention: Changes in attention z_{CS} ($z_{CS}, -1 > z_{CS} > +1$) to an active or predicted CS are proportional to the salience of the CS and are given by:

$$z_{CS} = in_{CS} ((K_4 OR (1 - |z_{CS}|) - K_5 (1 + z_{CS})))$$

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

$$OR = (Novelty_{total}^2 / (Novelty_{total}^2 + K_{11}^2))$$

$$z_{CS} > 0; \text{ when Novelty} > \text{Threshold}_{CS}$$

$$z_{CS} < 0; \text{ when Novelty} < \text{Threshold}_{CS}$$

$$\text{Threshold}_{CS} = K_5 / K_4$$

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

$$BUS = \Sigma X_{CS} V_{CS-US}$$

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:

$$\Delta V_{CS-US} = K_3 X_{CS} (\lambda_{US} - V_{CS-US})$$

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

$$X_{CS} \sim (\tau_{CS} + B_{CS}) z_{CS}$$

$$X_{CS} = K_7 (\ln_{CS} K_2 + (\ln_{CS} z_{CS}))$$

Pavlovian Conditioned Response: The US-specific CR is a sigmoid function of the total prediction of the US by all stimuli in the environment ($pre_{US} = pre_{US} + BUS$):

$$CR = K_6 \left(\frac{pre_{US}^2}{pre_{US}^2 + K_{12}^2} \right) (1 - OR)$$

Operant outcome: We assume that while the Pavlovian US takes a single value for each trial, the value of the operant outcome (R_{out}) is determined by the difference between two possible US outcomes (US), the US that occurs in the absence of the operant response ($US_{no\ act}$) versus the US that occurs following an operant response (US_{act}):

$$US = (US_{act} - US_{no\ act}) K_{13}$$

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, $US = US_{act} - US_{no\ act} = 1 - 0 = +1$ 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, $US = US_{act} - US_{no\ act} = 0 - (-1) = +1$ 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 (R_{out}) in a probabilistic fashion:

if $pre_{US} \leq 0.01$ then (80%) $R_{out} = US_{no\ act}$ (no response) vs. (20%) $R_{out} = US$ (response)

if $pre_{US} > 0.01$ then (20%) $R_{out} = US_{no\ act}$ (no response) vs. (80%) $R_{out} = US$ (response)

This way the $preUS$ and $Rout = US$ probability increases when $US > 0$ and decreases when $US < 0$.

For positive and negative reinforcement:

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

For positive and negative punishment:

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

The $Rout$ value was also used to calculate the prediction error ($\lambda US = Rout$), 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.

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 t-tests and one-way 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 see if these were significantly different from the critical z-score at p=0.05 level (1.645) using independent-t-tests. Alpha was 0.05 for all statistical analysis. All data were depicted as group mean \pm standard error of the mean (S.E.M.). 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, $\Delta F/F$ was z-scored on a trial-by-trial basis (10 second event-centered trials) and the mean value during the 2.5 seconds 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work was supported by NIH grants KL2TR002245 to M.G.K, DA042111 and DA048931 to E.S.C., 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.

REFERENCES

1. Schultz W, Dayan P & Montague PR A neural substrate of prediction and reward. *Science* (1997) doi:10.1126/science.275.5306.1593.
2. Dayan P & Balleine BW Reward, motivation, and reinforcement learning. *Neuron* (2002) doi:10.1016/S0896-6273(02)00963-7.
3. Dayan P & Niv Y Reinforcement learning: The Good, The Bad and The Ugly. *Current Opinion in Neurobiology* (2008) doi:10.1016/j.conb.2008.08.003.
4. Fiorillo CD, Newsome WT & Schultz W The temporal precision of reward prediction in dopamine neurons. *Nature Neuroscience* (2008) doi:10.1038/nn.2159.
5. Iordanova MD Dopaminergic modulation of appetitive and aversive predictive learning. *Reviews in the Neurosciences* (2009) doi:10.1515/REVNEURO.2009.20.5-6.383.
6. Flagel SB et al. A selective role for dopamine in stimulus-reward learning. *Nature* (2011) doi:10.1038/nature09588.
7. Glimcher PW Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* (2011) doi:10.1073/pnas.1014269108.
8. Steinberg EE et al. A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience* (2013) doi:10.1038/nn.3413.
9. Schultz W Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews Neuroscience* 17, 183 (2016). [PubMed: 26865020]
10. Rescorla R & Wagner A A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. in *Classical conditioning: current research and theory, Vol. 2* (1972).
11. Bayer HM & Glimcher PW Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* (2005) doi:10.1016/j.neuron.2005.05.020.
12. Lak A, Stauffer WR & Schultz W Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences of the United States of America* (2014) doi:10.1073/pnas.1321596111.
13. Eshel N, Tian J, Bukwich M & Uchida N Dopamine neurons share common response function for reward prediction error. *Nature neuroscience* 19, 479 (2016). [PubMed: 26854803]
14. Takahashi YK, Langdon AJ, Niv Y & Schoenbaum G Temporal specificity of reward prediction errors signaled by putative dopamine neurons in rat VTA depends on ventral striatum. *Neuron* 91, 182–193 (2016). [PubMed: 27292535]
15. Starkweather CK, Gershman SJ & Uchida N The Medial Prefrontal Cortex Shapes Dopamine Reward Prediction Errors under State Uncertainty. *Neuron* (2018) doi:10.1016/j.neuron.2018.03.036.
16. Phillips PEM, Stuber GD, Helen MLAV, Wightman RM & Carelli RM Subsecond dopamine release promotes cocaine seeking. *Nature* (2003) doi:10.1038/nature01476.
17. Day JJ & Carelli RM The nucleus accumbens and Pavlovian reward learning. *The Neuroscientist* 13, 148–159 (2007). [PubMed: 17404375]
18. Howe MW, Tierney PL, Sandberg SG, Phillips PEM & Graybiel AM Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* (2013) doi:10.1038/nature12475.
19. Hart AS, Rutledge RB, Glimcher PW & Phillips PEM Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *Journal of Neuroscience* (2014) doi:10.1523/JNEUROSCI.2489-13.2014.
20. Saddoris MP et al. Mesolimbic Dopamine Dynamically Tracks, and Is Causally Linked to, Discrete Aspects of Value-Based Decision Making. *Biological Psychiatry* (2015) doi:10.1016/j.biopsych.2014.10.024.
21. Bromberg-Martin ES, Matsumoto M & Hikosaka O Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. *Neuron* (2010) doi:10.1016/j.neuron.2010.11.022.

22. Howe MW & Dombeck DA Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature* (2016) doi:10.1038/nature18942.
23. Parker NF et al. Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nature Neuroscience* (2016) doi:10.1038/nn.4287.
24. Coddington LT & Dudman JT The timing of action determines reward prediction signals in identified midbrain dopamine neurons. *Nature Neuroscience* (2018) doi:10.1038/s41593-018-0245-7.
25. Da Silva JA, Tecuapetla F, Paixão V & Costa RM Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* (2018) doi:10.1038/nature25457.
26. Engelhard B et al. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* (2019) doi:10.1038/s41586-019-1261-9.
27. Redgrave P & Gurney K The short-latency dopamine signal: a role in discovering novel actions? *Nature reviews*. 7, 967–975 (2006).
28. Ungless MA Dopamine: the salient issue. *Trends in neurosciences*. 27, 702–706 (2004). [PubMed: 15541509]
29. Brischoux F, Chakraborty S, Brierley DI & Ungless MA Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proceedings of the National Academy of Sciences* (2009) doi:10.1073/pnas.0811507106.
30. Valenti O, Lodge DJ & Grace AA Aversive stimuli alter ventral tegmental area dopamine neuron activity via a common action in the ventral hippocampus. *The journal of neuroscience : the official journal of the Society for Neuroscience*. 31, 4280–4289 (2011). [PubMed: 21411669]
31. McCutcheon JE, Ebner SR, Loriaux AL & Roitman MF Encoding of Aversion by Dopamine and the Nucleus Accumbens. *Frontiers in neuroscience*. 6, 137 (2012). [PubMed: 23055953]
32. Patriarchi T et al. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science* (2018).
33. Lewis AS, Calipari ES & Siciliano CA Toward Standardized Guidelines for Investigating Neural Circuit Control of Behavior in Animal Research. *eNeuro* 8, (2021).
34. Kutlu MG et al. A novel multidimensional reinforcement task in mice elucidates sex-specific behavioral strategies. *Neuropsychopharmacology* (2020) doi:10.1038/s41386-020-0692-1.
35. Oleson EB, Gentry RN, Chioma VC & Cheer JF Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *Journal of Neuroscience* (2012) doi:10.1523/JNEUROSCI.3087-12.2012.
36. Christianson JP et al. Inhibition of fear by learned safety signals: A mini-symposium review. *Journal of Neuroscience* (2012) doi:10.1523/JNEUROSCI.3340-12.2012.
37. Menegas W, Akiti K, Amo R, Uchida N & Watabe-Uchida M Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nature Neuroscience* (2018) doi:10.1038/s41593-018-0222-1.
38. Wenzel JM et al. Phasic Dopamine Signals in the Nucleus Accumbens that Cause Active Avoidance Require Endocannabinoid Mobilization in the Midbrain. *Current Biology* (2018) doi:10.1016/j.cub.2018.03.037.
39. de Jong JW et al. A Neural Circuit Mechanism for Encoding Aversive Stimuli in the Mesolimbic Dopamine System. *Neuron* (2019) doi:10.1016/j.neuron.2018.11.005.
40. Stelly CE et al. Pattern of dopamine signaling during aversive events predicts active avoidance learning. *Proceedings of the National Academy of Sciences of the United States of America* (2019) doi:10.1073/pnas.1904249116.
41. Pavlov IP Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. (Oxford Univ. Press, 1927).
42. Pearce JM A model for stimulus generalization in Pavlovian conditioning. *The psychological review*. 94, 61–73 (1987). [PubMed: 3823305]
43. Brandon SE, Vogel EH & Wagner AR A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behavioural brain research*. 110, 67–72 (2000). [PubMed: 10802304]

44. Sokolov EN The orienting response, and future directions of its development. *The Pavlovian Journal of Biological Science* (1990) doi:10.1007/BF02974268.
45. Sokolov EN Neuronal models and the orienting reflex. in *The central nervous system and behavior*, 3rd conference 187–276 (Josiah Macy Jr Foundation, 1960).
46. Schmajuk NA, Lam Y-W & Gray JA Latent inhibition: A neural network approach. *Journal of Experimental Psychology: Animal Behavior Processes* 22, 321 (1996). [PubMed: 8691162]
47. Lubow RE & Moore AU Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of comparative and physiological psychology* 52, 415 (1959). [PubMed: 14418647]
48. Channell S & Hall G Facilitation and retardation of discrimination learning after exposure to the stimuli. *Journal of experimental psychology*. 7, 437–446 (1981). [PubMed: 7288372]
49. Lubow RE Latent inhibition as a measure of learned inattention: Some problems and solutions. In *Behavioural Brain Research* (1997). doi:10.1016/S0166-4328(97)02307-3.
50. Quintero E et al. Effects of context novelty vs. familiarity on latent inhibition with a conditioned taste aversion procedure. *Behavioural Processes* (2011) doi:10.1016/j.beproc.2010.12.011.
51. Sutton RS & Barto AG Time-Derivative Models of Pavlovian Reinforcement. *Learning and Computational Neuroscience: Foundations of Adaptive Networks* (1990) doi:10.1111/j.1748-1716.1960.tb01900.x.
52. Young AMJ, Joseph MH & Gray JA Latent inhibition of conditioned dopamine release in rat nucleus accumbens. *Neuroscience* 54, 5–9 (1993). [PubMed: 8515846]
53. Sharpe MJ et al. Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience* (2017) doi:10.1038/nn.4538.
54. Kutlu MG & Schmajuk NA Solving Pavlov's puzzle: Attentional, associative, and flexible configural mechanisms in classical conditioning. *Learning and Behavior* (2012) doi:10.3758/s13420-012-0083-5.
55. Kamin LJ Predictability, surprise, attention and conditioning. *Punishment and aversive behavior* (1969) doi:10.2105/AJPH.2006.090910.
56. Holland PC Overshadowing and blocking as acquisition deficits: No recovery after extinction of overshadowing or blocking cues. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology* (1999) doi:10.1080/027249999393022.
57. Parker KE et al. A Paranigral VTA Nociceptin Circuit that Constrains Motivation for Reward. *Cell*. 178, 653–671.e19 (2019). [PubMed: 31348890]
58. Young AMJ Increased extracellular dopamine in nucleus accumbens in response to unconditioned and conditioned aversive stimuli: Studies using 1 min microdialysis in rats. *Journal of Neuroscience Methods* (2004) doi:10.1016/j.jneumeth.2004.03.003.
59. Mileykovskiy B & Morales M Duration of inhibition of ventral tegmental area dopamine neurons encodes a level of conditioned fear. *Journal of Neuroscience* (2011) doi:10.1523/JNEUROSCI.5731-10.2011.
60. Budygin EA et al. Aversive stimulus differentially triggers subsecond dopamine release in reward regions. *Neuroscience* (2012) doi:10.1016/j.neuroscience.2011.10.056.
61. Dickinson A, Hall G & Mackintosh NJ Surprise and the attenuation of blocking. *Journal of Experimental Psychology: Animal Behavior Processes* (1976) doi:10.1037/0097-7403.2.4.313.
62. Holland PC Pretraining a compound conditioned stimulus reduces unblocking. *Bulletin of the Psychonomic Society* (1985) doi:10.3758/BF03329837.
63. Mackintosh NJ & Turner C Blocking as a function of novelty of CS and predictability of UCS. *The Quarterly journal of experimental psychology* (1971) doi:10.1080/14640747108400245.
64. Mackintosh NJ A theory of attention: variations in the associability of stimuli with reinforcement. *Psychological review* 82, 276 (1975).
65. Pearce JM & Hall G A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review* 87, 532 (1980). [PubMed: 7443916]
66. Comoli E et al. A direct projection from superior colliculus to substantia nigra for detecting salient visual events. *Nature Neuroscience* (2003) doi:10.1038/nn1113.

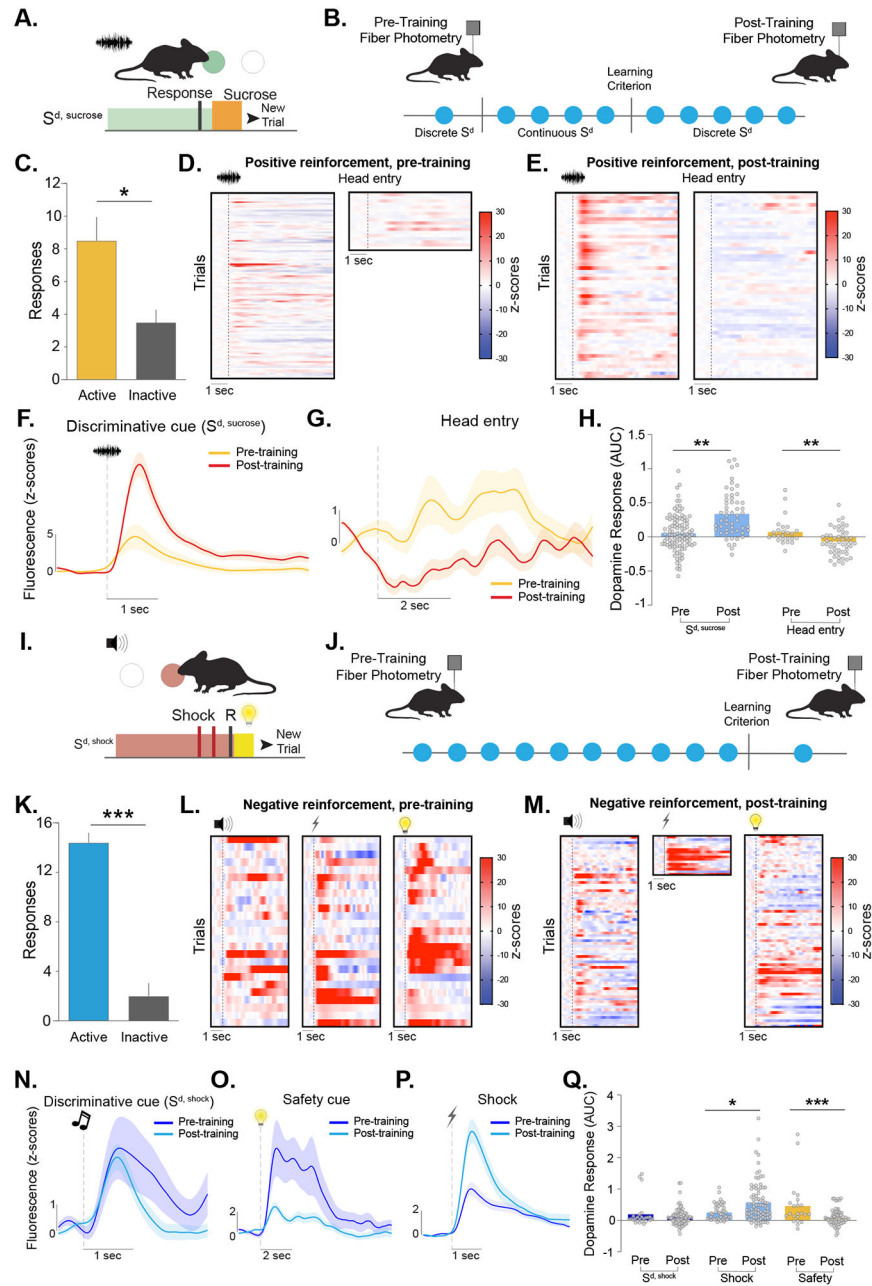


Figure 1. Dopamine release patterns in the nucleus accumbens core do not track reward prediction error in negative reinforcement tasks.

(A) Positive reinforcement task design. Mice were trained to nose poke during a discriminative cue ($S^{d,sucrose}$) to obtain sucrose. (B) Fiber photometry was performed using the fluorescent dopamine sensor dLight1.1 during the first (pre-training) and last (post-training) session of positive reinforcement. (C) Mice made more active than inactive nose pokes during the post-training session (paired t-test, $t_6=3.18$, $p=0.024$; $n=6$). Heatmaps showing dopamine responses aligned around $S^{d,sucrose}$ and head entry during (D) pre-training and (E) post-training. Each row represents a single presentation of $S^{d,sucrose}$ and a single head entry response for a given trial. Averaged dopamine traces showing (F) $S^{d,sucrose}$

and **(G)** head entry responses during pre-training and post-training sessions. **(H)** Dopamine responses to the $S^{d,sucrose}$ increased over training (Nested ANOVA, $F_{(1,153)}=10.79$, $p=0.0013$). Dopamine responses following head entries into the sucrose port decreased with training (Nested ANOVA, $F_{(1,75)}=11.17$, $p=0.0013$). **(I)** Negative reinforcement task design. Mice learned to respond during a separate discriminative cue ($S^{d,shock}$) to avoid a series of footshocks. A safety cue was presented at the end of each trial **(J)** Experimental design for recording dopamine responses during the task. **(K)** Behavioral performance during the post-training recording session (paired t-test, $t_4=9.35$, $p<0.001$; $n=5$). Heatmaps of dopamine responses aligned around $S^{d,shock}$, footshock, and safety cue during **(L)** pre-training and **(M)** post-training sessions. Each row represents a single presentation of $S^{d,shock}$, safety cue, and post-training footshock as well as the first presentation of pre-trainings socks for a given trial. Averaged dopamine traces showing **(N)** $S^{d,shock}$, **(O)** footshock, and **(P)** safety cue responses pre-training and post-training. **(Q)** Dopamine response did not change to the $S^{d,shock}$ (Nested ANOVA, $F_{(1,96)}=1.52$, $p=0.220$). Dopamine release to the footshock was increased (Nested ANOVA, $F_{(1,127)}=4.00$, $p=0.047$), and dopamine responses to the safety cue was reduced (Nested ANOVA, $F_{(1,95)}=15.46$, $p=0.0002$). Data represented as mean \pm S.E.M. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

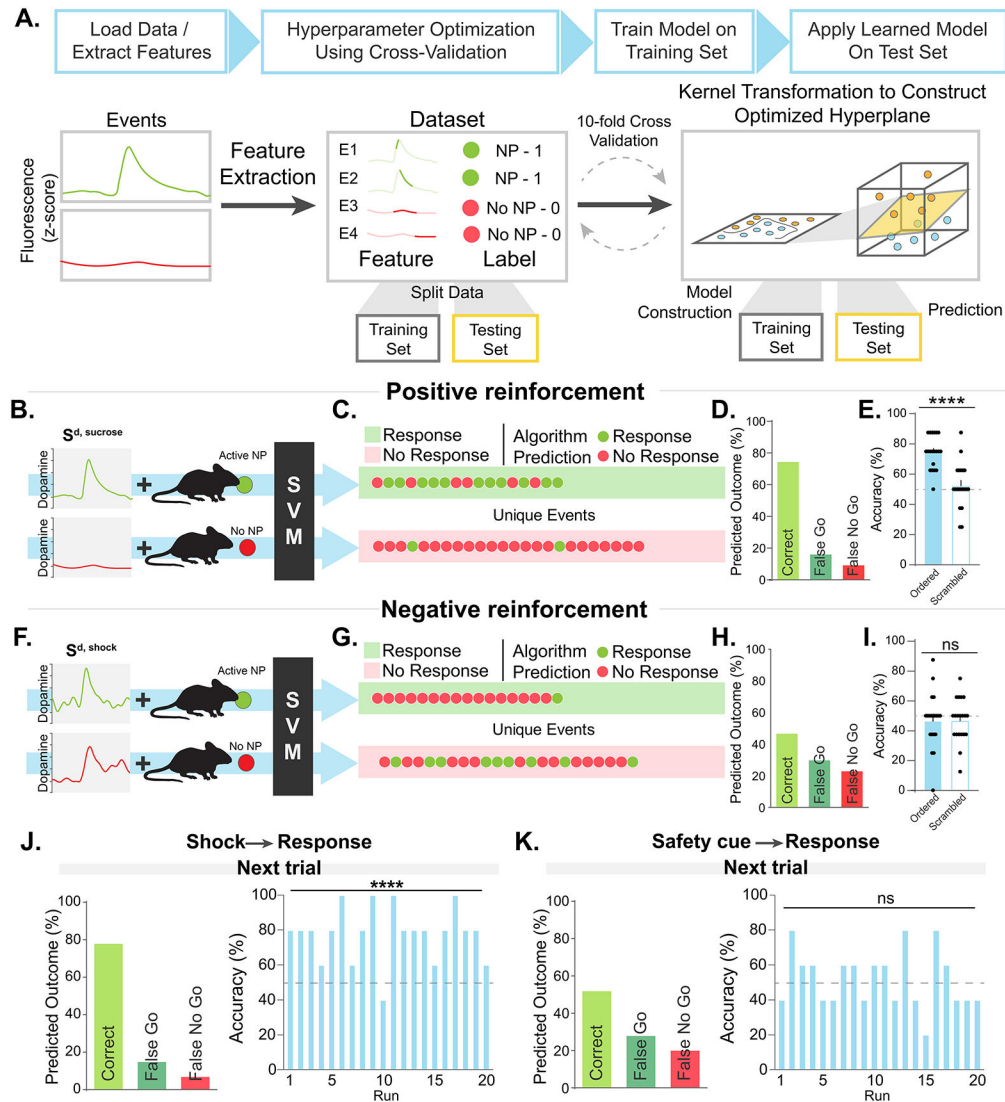


Figure 2. Dopamine responses to footshocks, but not safety cues, predict future behavior during negative reinforcement.

(A) Supervised machine learning was used to define whether dopamine signals around each behavioral event predicted current and/or future trial behavior. Support vector machine (SVM) algorithms were used. Datasets from dopamine recordings were coded depending on whether animals made a response or not in each trial. Data were then split into a training set and a testing set where the training set was used to construct an optimized hyperplane for behavior prediction and the testing set was used to test the accuracy of those predictions in an iterative fashion. (B-E) The SVM algorithm accurately predicted the current trial behavioral response based on the dopamine response to the positive reinforcement discriminative cue ($S^{d, sucrose}$; >75% accuracy; unpaired t-test ordered vs scrambled controls, $t_{38}=5.34$, $p<0.0001$; $N=69$ trials) with few errors (~20% false go and false no go predictions; opposite dot color). (F-I) For negative reinforcement, the algorithm was unable to use the $S^{d, shock}$ dopamine response to predict whether animals would or would not respond in the current trial (~45% accuracy; unpaired t-test ordered vs scrambled

controls, $t_{38}=0.000$, $p>0.99$; $N=69$ trials). **(J)** During negative reinforcement, when animals did not respond during the $S^{d,shock}$ they received a footshock. The SVM algorithm was able to predict whether an animal would nose-poke during the $S^{d,shock}$ or not on the next trial based on the dopamine response to the shock itself (independent t-test, $t_{19}=7.95$, $p<0.0001$; $N=30$ trials; **see Figure S3 for additional analyses**). **(K)** The SVM algorithm was unable to predict behavior in the subsequent trial based on the dopamine response to the safety cue on trials when animals did not respond correctly during $S^{d,shock}$, suggesting that dopamine responses to the safety cue are not error-based learning signals ($>75\%$ accuracy; independent t-test, $t_{19}=0.55$, $p=0.59$; $\sim 50\%$ accuracy; $N=30$ trials; **see Figure S3 for additional analyses**). Data represented as mean \pm S.E.M., **** $p < 0.0001$; ns, not significant.

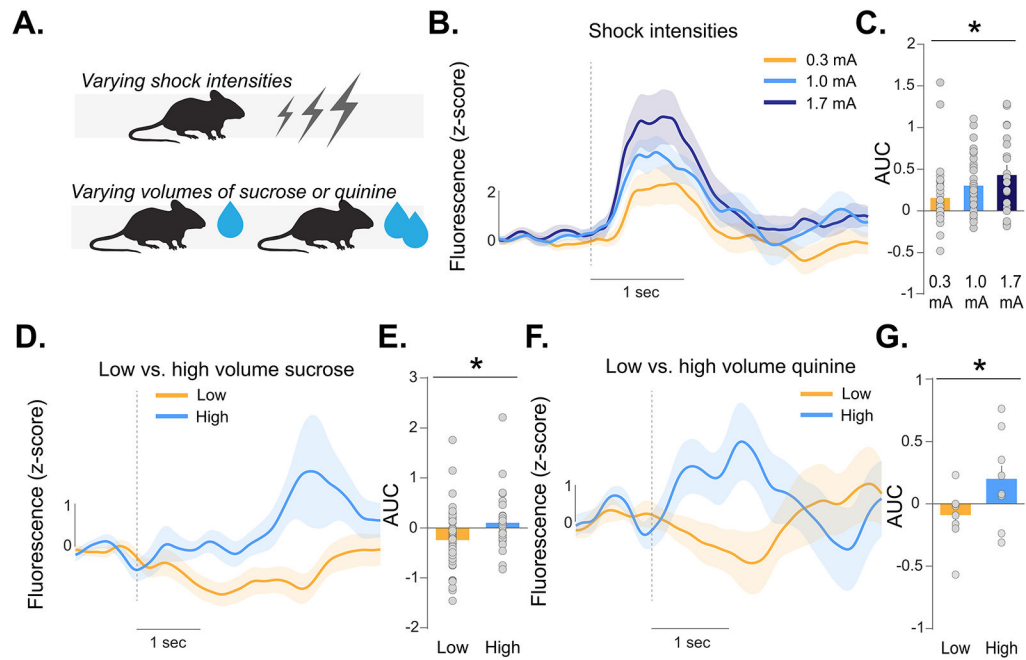


Figure 3. Accumbal dopamine release signals stimulus saliency but not stimulus valence. (A) A series of experiments were run to dissociate responses to saliency and valence. Mice received varying intensities of footshocks or varying volumes of sucrose or quinine. (B-C) Dopamine response in the NAc core increased with increasing shock intensities (Nested ANOVA, $F_{2,83}=4.16$, $p=0.019$; 0.3mA vs. 1.7mA $p=0.008$; $n=5$). (D-E) Increasing the volume of both sucrose (Nested ANOVA, $F_{(1,85)}=6.31$, $p=0.014$; $n=3$) or (F-G) the bitter tastant quinine resulted in increased dopamine responses to licks (Nested ANOVA, $F_{(1,18)}=4.60$, $p=0.048$; $n=3$). Data represented as mean \pm S.E.M. * $p < 0.05$.

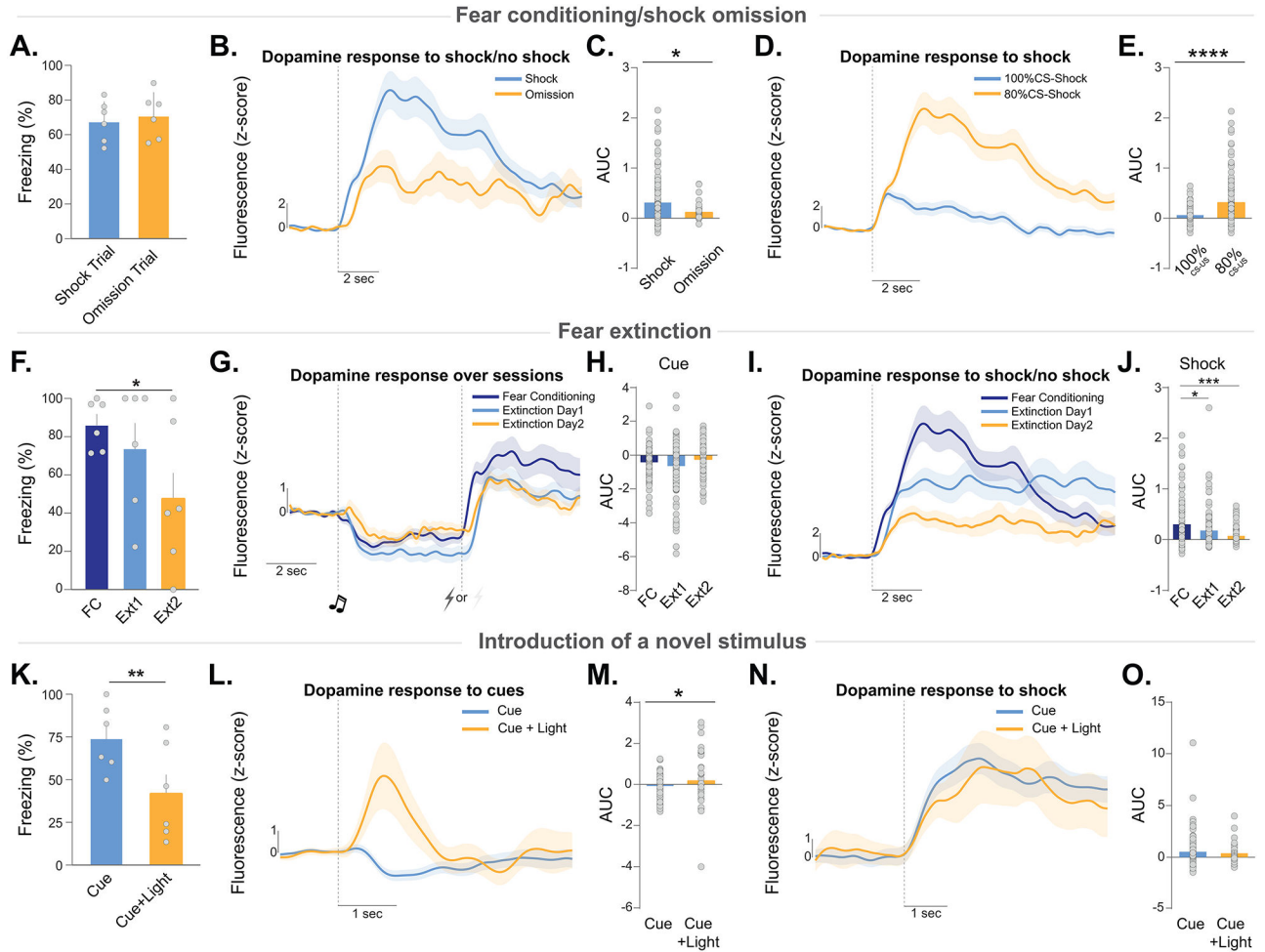


Figure 4. Dopamine release in the nucleus accumbens core does not follow prediction error patterns during unexpected addition and omission of stimuli.

A series of experiments were run to test how unexpected addition and omission of predicted and unpredicted stimuli influence dopamine release patterns. Mice first were trained with Pavlovian fear conditioning. **(A)** Footshocks were unexpectedly omitted in 20% of trials. Freezing during each trial type was not different (paired t-test, $t_5=1.47$, $p=0.20$; $n=6$). **(B-C)** Dopamine responses at the time of the omitted footshocks were larger than the critical z-score at $p=0.05$ level, indicating a significant dopamine event even though the shock itself was not presented (independent t-test; $t_{29}=3.92$, $p=0.0005$). Dopamine responses at the time of the predicted footshock were lower in amplitude than when the footshock was present (Nested ANOVA, $F_{(1,156)}=7.39$ $p=0.0073$). **(D-E)** Dopamine responses to the footshocks during the omission session (when shock probability was 80%) were larger than the shock response during the initial fear conditioning session (when shock probability was 100%) (Nested ANOVA, $F_{(1,222)}=21.48$, $p<0.0001$). **(F)** Freezing during acquisition and extinction (Fear conditioning vs. Extinction 2, paired t-test, $t_5=3.31$, $p=0.021$; $n=6$). **(G-H)** Dopamine response to the cue did not change over extinction, although it returned to baseline (Nested ANOVA, $F_{(2,331)}=3.37$, $p=0.035$; Early extinction vs. Late extinction, $p=0.009$). **(I-J)** Dopamine responses to the shock delivery period decreased over extinction.

Responses to omitted shocks during extinction were smaller than when footshocks were present. (Nested ANOVA, $F_{2,329}=9.76$, $p=0.0001$; Fear conditioning vs. Extinction Day 1 $p=0.0005$; Fear conditioning vs. Extinction Day 2 $p<0.0001$). **(K)** A novel, neutral light was presented in 20% of trials. Freezing to the tone and tone + light trials show that the novel cue reduced freezing (unpaired t-test, $t_5=4.23$, $p=0.008$; $n=6$) and **(L-M)** dopamine responses to the tone + novel cue was increased (Nested ANOVA, $F_{1,156}=6.44$, $p=0.012$). **(N-O)** Dopamine response to the footshocks did not differ between trials (Nested ANOVA, $F_{1,156}=0.35$, $p=0.557$). Data represented as mean \pm S.E.M. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

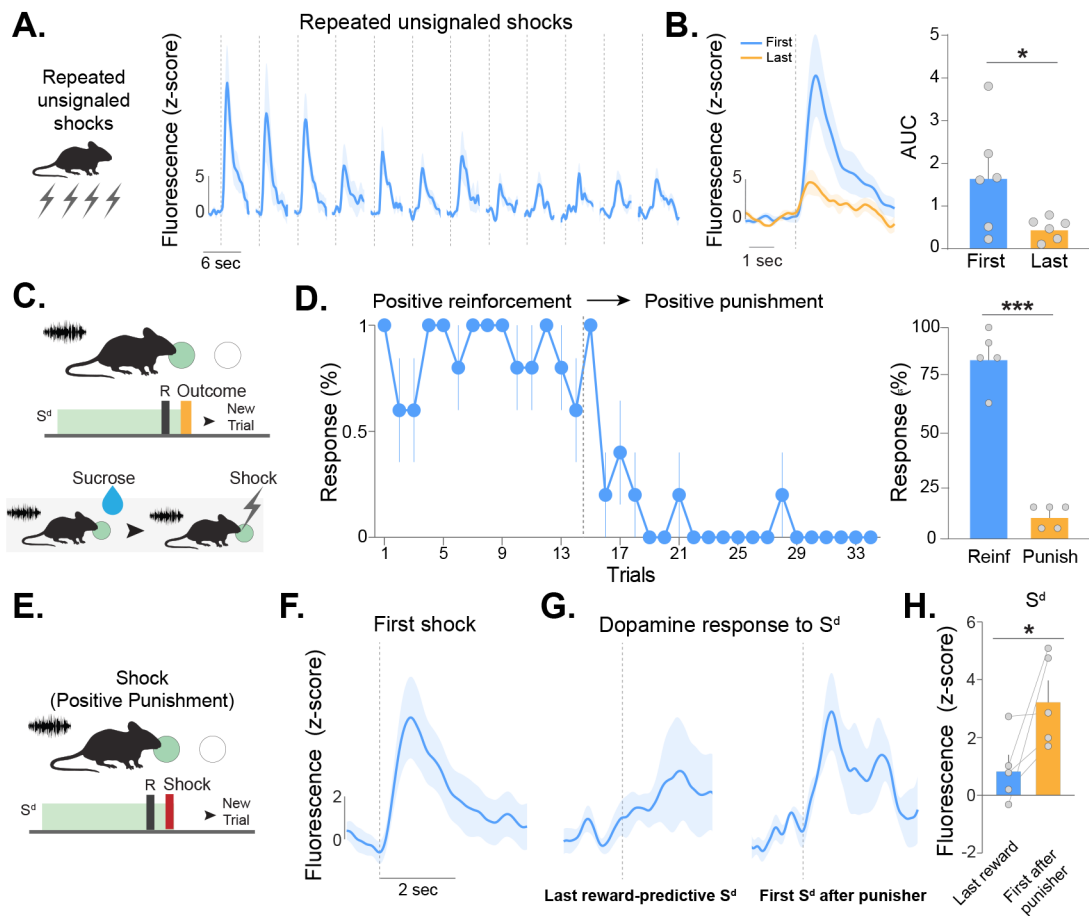


Figure 5. NAc core dopamine responses decrease with repeated presentations of stimuli and increase during worse than expected outcomes.

(A) Mice received repeated presentations of footshocks with constant intensity on a fixed interval schedule. (B) Dopamine responses to footshocks decreased over time (paired t-test, $t_5=2.60$, $p=0.047$; $n=6$). (C) Mice were trained to nose poke during a discriminative cue (S^d) for sucrose. In subsequent trials, responses during the same S^d now resulted in a single footshock (i.e., punishment) (D) Switching from positive reinforcement to punishment (denoted by dotted line) decreased operant responses. Nose pokes were higher during the positive reinforcement phase compared to the punishment phase (paired t-test, $t_4=11.57$, $p=0.0003$; $n=5$, 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, i.e. a negative prediction error. (F) The dopamine response to the first footshock following the operant response was positive – even though it was an unexpected negative outcome. (G-H) The dopamine responses to the S^d also increased, even though it represented a worse outcome than the previous association for a cue with which the animal had extensive previous experience (paired t-test, $t_4=2.55$, $p=0.031$). Data represented as mean \pm S.E.M. * $p < 0.05$, *** $p < 0.001$.

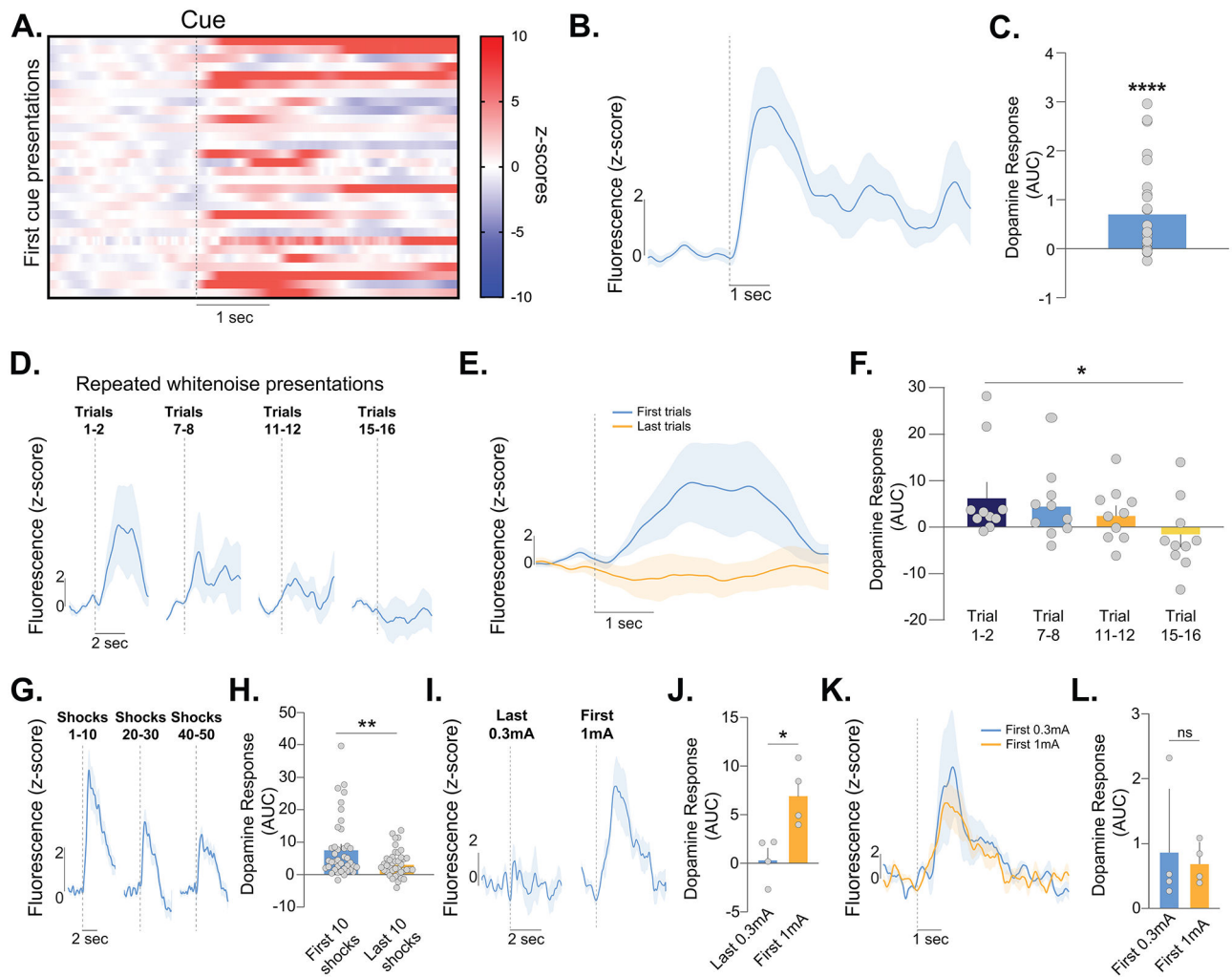


Figure 6: Neutral stimuli evoke dopamine release in the nucleus accumbens core.

(A) Heatmap showing the dopamine response to the first-time presentations of cues in this study. Each presentation is from the positive and negative reinforcement training on the first trial across all animals, before they had learned the task or made any behavioral responses, thus these stimuli did not have associative value. (B-C) Average dopamine response to discriminative stimuli before acquiring value in each reinforcement task (independent t-test, $t_{25}=4.08$, $p=0.0004$; $n=10$). (D) A separate experiment was run where white noise was presented 16 times and was not associated with an unconditioned stimulus. Presentation of the neutral white noise evoked a positive dopamine response, which decreased with repeated presentations. (E-F) Average dopamine response to the white noise decreased between the first versus last presentations (RM ANOVA, $F_{(2,594,23,35)}=3.20$, $p=0.048$; Dunnett multiple comparison to Trial1-2 vs. Trial 15-16, $p=0.036$; $n=5$). (G-H) Repeated presentations of a mild footshock (0.3mA) decreased dopamine response to footshocks as the novelty decreased (paired t-test, $t_{39}=3.455$, $p=0.0013$). (I-J) Increasing intensity of the footshock reinstated the habituated dopamine response to the footshock (last 0.3mA vs first 1mA footshock, paired t-test, $t_3=4.94$, $p=0.0159$). (K-L) When compared to the first presentation of the low intensity (0.3mA) shock presentation, the first presentation

of the high intensity (1mA) footshock presentation did not yield a stronger dopamine response after habituation (paired t-test, $t_3=0.470$, $p=0.67$; $n=4$). Together these data show that dopamine signals perceived saliency where it is influenced by novelty and intensity of stimuli in the environment regardless of valence. Data represented as mean \pm S.E.M. * $p < 0.05$, ** $p < 0.05$, **** $p < 0.0001$, ns = not significant.

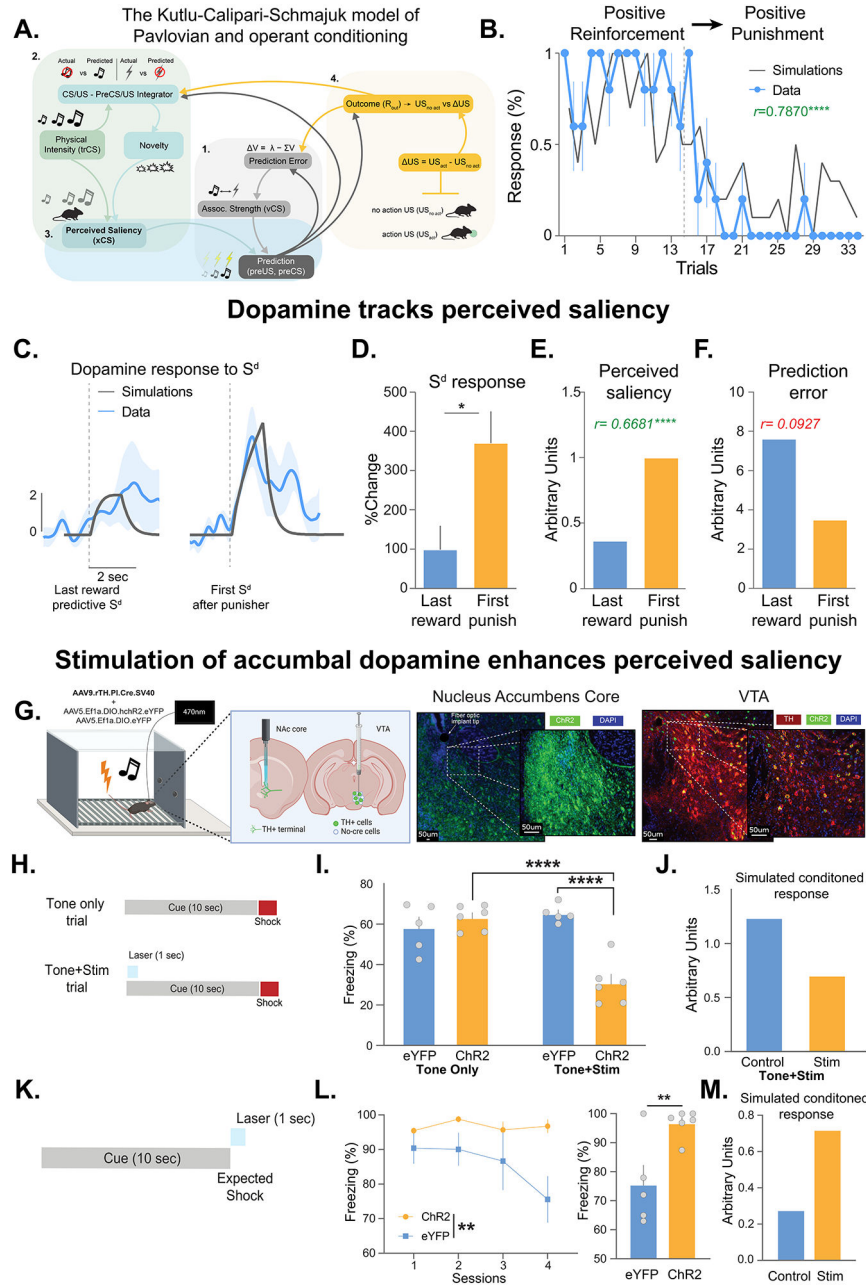


Figure 7: Dopamine release in the nucleus accumbens core tracks perceived saliency. (A) The Kutlu-Calipari-Schmajuk (KCS) model. The model has 4 core components. 1) *Associative component*: Based on a Rescorla-Wager type prediction error term. 2) *Attentional component*: Mismatch between predicted/unpredicted 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 a prediction of an outcome. For operant responses, the value of an outcome is calculated as the difference between the unconditioned stimulus value before and after the operant response and predicts future responding in a probabilistic fashion. (B-F)

Experiments from Figure 4 were replotted to map onto model simulations. Contingency switch from positive reinforcement to punishment represents a worse than expected outcome – i.e. a negative prediction error. **(B)** Model simulations of reinforcement behavior (grey line) overlaid with experimental data (blue line). 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 (grey), and dopamine responses to (blue), the cue increased. **(D)** The dopamine response to the cue after the first punisher was increased (paired t-test, $t_4=2.76$, $p=0.025$). KCS model simulations show that perceived saliency **(E)**, $R=0.67$, $p<0.0001$) matches dopamine response patterns during the contingency switch but prediction error does not **(F)**, $r=0.092$, $p=0.2264$). **(G-M)** Optogenetics studies were designed to test whether behavioral responses change in response to increasing dopamine as predicted by changes in perceived saliency. **(G)** AAV.TH.Cre and AAV.DIO.ChR2 or eYFP were injected into the VTA to achieve specific expression of opsins in dopamine neurons. A fiberoptic was placed above the NAc core to stimulate dopamine release from terminals selectively in the NAc core. **(H)** A fear conditioning experiment was run where dopamine release was evoked during the cue on 25% of cue-shock pairings. **(I)** Increasing NAc core dopamine decreased freezing in the ChR2 group compared to eYFP controls (2 way-ANOVA, trial type x group interaction, $F_{(1,20)}=17.84$, $p=0.0004$; Sidak' multiple comparison ChR2-Tone+Stim vs. eYFP-Tone+Stim, $p<0.0001$; ChR2-Tone+Stim vs. ChR2-Tone only $p<0.0001$; $n=5-6$) and tone only trials in the same animals (Sidak' multiple comparison ChR2-Tone+Stim vs. ChR2-Tone only, $p=0.0002$). **(J)** Simulations from the KCS model show that this behavioral response is predicted by increased perceived saliency, but not other prediction-based parameters. **(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 (RM ANOVA, Group main effect, $F_{1,9}=5.90$, $p=0.038$; Last 4 trial block, unpaired t-test, $t_9=3.32$, $p=0.0089$; $n=5-6$). **(M)** KCS model simulations show enhancing perceived saliency of the omitted shocks prevents extinction of the conditioned response. Data represented as mean \pm S.E.M. * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$.