Extinction and Reinstatement of Phasic Dopamine Signals in the Nucleus Accumbens Core during Pavlovian Conditioning

Ceyhun Sunsay and
Department of Psychology, 3400 Broadway, Indiana University Northwest, Gary, IN 46408

George V. Rebec
Program in Neuroscience, Department of Psychological and Brain Sciences, 1101 E. Tenth Street, Indiana University, Bloomington, IN 47405

Abstract

The prediction-error model of dopamine (DA) signaling has largely been confirmed with various appetitive Pavlovian conditioning procedures and has been supported in tests of Pavlovian extinction. Studies have repeatedly shown, however, that extinction does not erase the original memory of conditioning as the prediction-error model presumes, putting the model at odds with contemporary views that treat extinction as an episode of learning rather than unlearning of conditioning. Here, we combined fast-scan cyclic voltammetry (FSCV) with appetitive Pavlovian conditioning to assess DA release directly during extinction and reinstatement. DA was monitored in the nucleus accumbens core (NAc), which plays a key role in reward processing. Following at least 4 daily sessions of 16 tone-food pairings, FSCV was performed while rats received additional tone-food pairings followed by tone alone presentations (i.e., extinction). Acquisition memory was reinstated with non-contingent presentations of reward and then tested with cue presentation. Tone-food pairings produced transient (1–3 s) DA release in response to tone. During extinction, the amplitude of the DA response decreased significantly. Following presentation of two non-contingent food pellets, subsequent tone presentation reinstated the DA signal. Our results support the prediction-error model for appetitive Pavlovian extinction but not for reinstatement.

Keywords

Dopamine (DA) has been implicated in various aspects of reward processing such as consumption (Wise, 1978, 1982), incentive value (Flagel et al, 2011), anticipation (Schultz, 1998) and the selection of appropriate action-outcome routines (Redgrave & Gurney, 2006). Studies of DA involvement in reward anticipation are largely based on single-unit recording and measurements of DA release conducted in conjunction with operant and Pavlovian conditioning. The data indicate that first exposure to an appetitive unconditional stimulus (US) (i.e., reward) triggers increased firing in presumed mid-brain DA neurons with no

Correspondence: George V. Rebec, Psychology Building, Indiana University, 1101 E. Tenth Street, Bloomington, IN 47405-7007; Phone: 812-855-4832; Fax: 812-855-4520; rebec@indiana.edu.

discernible change in firing to presentation of a conditional stimulus (CS) that precedes the reward. Importantly, as conditioning proceeds, neuronal firing decreases with the presentation of reward, while the neuronal firing to the presentation of the CS increases. This shift is argued to reflect anticipation of reward upon perception of the CS (Mirenowicz & Schultz, 1994; 1996). The subsequent use of fast-scan cyclic voltammetry (FSCV) confirmed that the shift in neuronal activity paralleled the release of DA (Day, Stuber, Wightman & Carelli, 2007; Stuber et al., 2008; Sunsay & Rebec, 2008).

Studies also showed that DA neuronal firing is depressed when an anticipated reward is no longer delivered in extinction and conditioned inhibition procedures (Pan, Schmidt, Wickens & Hyland, 2008; Tobler, Dickinson & Schultz, 2003). FSCV studies using operant intra-cranial self-stimulation and cocaine self-administration showed decreases in DA amplitudes during extinction (Owesson-White, Cheer, Beyene, Carelli & Wightman, 2008; Stuber, Wightman & Carelli, 2005), suggesting that DA release parallels firing rate of DA neurons in coding only positive prediction errors. A more recent study, using a probabilistic decision-making task imposed on operant conditioning controlling for reward variance and thus the degree of uncertainty between choosing the correct lever, showed equivalence in the amount of DA release in the NAc when animals received less reward than they expected compared to more reward than they expected (Hart, Rutledge, Glimcher & Phillips, 2014). In effect, positive and negative prediction errors were represented symmetrically with DA concentration in the NAc. Together, these findings support the prediction-error hypothesis (Schultz, Dayan & Montague, 1997), which states that changes in DA activity associated with reward anticipation are correlated with the amount of CS-US pairing. Recent evidence also suggests that the mesolimbic DA pathway may be involved in extinction (Hart et. al., 2014).

Although DA release appears to be a neurochemical index of associative learning, some key aspects of the prediction-error hypothesis are not entirely consistent with the literature on animal Pavlovian conditioning. The prediction-error model, like the Rescorla and Wagner model (1972), treats extinction as an unlearning episode. According to both models, extinction erases the CS-US association that was acquired during conditioning. Ample evidence now indicates, however, that the unlearning explanation for extinction is invalid. The data suggest instead that extinction involves new learning, such as a CS-no US association (Bouton, 1993). For example, an extinguished CS-US association is renewed upon return to the physical context in which the original learning took place, while extinction performance is still remembered upon return to the extinction context (Bouton & Bolles, 1979a). Extinction as new learning is supported by reinstatement experiments in which non-contingent presentations of a reinforcer reinstate the association that was previously extinguished (Bouton & Bolles, 1979b). Thus, what is learned in extinction is not lost; it is remembered as another association. Collectively, these studies emphasize the importance of memory retrieval in Pavlovian conditioning (Bouton, 1993). The memory of a CS-US association can be triggered by the correct cue such as the context or the passage of time as in the case of the spontaneous recovery effect (Rescorla, 2004).

Electrophysiological and electrochemical studies of DA signaling produced results inconsistent with the standard models of DA based on prediction errors. Following

extinction, the amplitude of the presumed DA signal was shown to recover with the passage of time (Pan et. al., 2008). Standard prediction-error models cannot accommodate such recovery as they are solely based on updating the weights that code the strength of the CS-US association. Standard prediction-error models are also unable to deal with the recovery of the conditioned response following a context switch or non-contingent reward delivery. FSCV studies, done in conjunction with operant conditioning and intra-cranial self-stimulation as reward, showed such recovery of DA amplitudes in a reinstatement procedure (Owesson-White, et. al., 2008; Stuber, et. al., 2005).

Recent reformulations of prediction-error models have attempted to address these weaknesses. Pan et al (2008) solve the spontaneous recovery problem by incorporating an additional weight that represents the strength of inhibition acquired during extinction and different rates of forgetting for conditioning and extinction weights. Therefore, the weight of inhibitory extinction memory decays faster during the retention interval and contributes to the retrieval of the conditioning memory (i.e., the spontaneous recovery effect). Electrophysiological recordings from DA neurons support this model (Pan et al., 2008).

The noteworthy difficulty, however, is the inability of the model to explain post-extinction recovery effects that occur with context change after extinction (i.e., the renewal effect) and with non-contingent reward presentations (i.e., the reinstatement effect). Two recent models are able to account for such post-extinction recovery effects since physical properties of events such as the context are active factors in either categorizing events or inferring latent causes of events. One of these models solves post-extinction recovery effects by operating the prediction error on Gaussian multivariate states (Redish, Jensen, Johnson & Kurth-Nelson, 2007). According to this model, animals are equipped with mechanisms that enable them to identify factors such as a new context that co-occur with new changes such as extinction. When extinction takes place in a new context, the omission of expected rewards generates a negative prediction error presumably driven by tonic DA signals, directing the animal to increase its attention to identify the factor or factors, such as new context, that co-occur with the new outcome (i.e., no reward). The model assumes that the omission of the expected reward consequently splits the state space, which refers to the current situation the animal experiences in the traditional temporal difference models. Because of this categorization process, the animal is able to represent acquisition and extinction in two different categories. Therefore, when reintroduced to the context in which acquisition had previously taken place, because of the split space, the animal "retrieves" the memory of conditioning; when reintroduced to the extinction context, the animal "retrieves" the extinction memory. The strength of the model is the active role given to negative prediction error that not only weakens the CS-US association but also motivates the animal by searching for the cause of the new outcome. Another strength of the model is its ability to partition the state space allowing varieties of inputs such as physical or internal context to serve as a cue that could be involved in retrieving the right association at the time of test. The model treats non-contingent rewards as states that can signal the original conditioning state similar to contexts. Therefore, the model is able to explain the reinstatement effect unlike the tripartite mechanism presented by Pan et al (2008). However, the model suffers from other problems, such as the context-specificity of latent inhibition, which does not involve negative prediction error (i.e., the retardation of conditioning of a cue following its

pre-exposure) and other forms of the renewal effect such as ABC renewal where the conditioned response is acquired in context A, extinguished in context B but renewed in context C.

A Bayesian model of learning solves these difficulties because it is not based on negative prediction as it does not assume association between cue and reward (Gershman, Blei & Niv, 2010). According to this model, animals observe events and outcomes and predict their causes that are hidden to them (i.e., latent causes). When observations differ from previous ones, animals start updating their previous beliefs about the latent cause. The authors posit a particle-filtering process that updates previous beliefs and generates a new latent cause. When observations are consistent with and similar to each other, such as cues being followed by the same reward consistently, the particle-filtering process clusters all observations together, averages them, and generates a single latent cause such as the acquisition phase. In a sense, the animal infers that acquisition is being carried out. Consequently, in a subsequent trial, the animal predicts the reward based on the Bayes rule. If a different outcome is observed such as omission of reward, the particle-filtering process generates a new cluster. As similar new observations accumulate, the new cluster is segregated by the previous cluster that represents acquisition, and the animal infers a new latent cause (i.e., extinction phase). When acquisition and extinction take place in different contexts, the return to the original context in which conditioning had previously taken place enables the animal to infer that the first latent cause is active and consequently the animal predicts reward. Although the model is able to account for the variations of the renewal effect as well as the context specificity of the latent-inhibition effect, it is unable to explain the spontaneous-recovery effect since there is no variable in the model that represents the dynamics of time. The model, however, explains the reinstatement effect. Put simply, the animal infers that the acquisition phase has returned with the presentation of non-contingent rewards and predicts the reward when the cue is turned on following the non-contingent reward.

Given the variations between these models in their treatments of the reinstatement effect, it is important to show whether reinstatement of DA release occurs. FSCV carried out in conjunction with operant conditioning procedures showed reinstatement of DA release following extinction (Owesson-White, et. al., 2008; Stuber, et. al., 2005). Operant conditioning involves learning the consequences of responses rather than antecedents of consequences as in the case of Pavlovian conditioning. Here, we used a Pavlovian conditioning procedure in conjunction with FSCV to investigate DA release in response to a CS during behavioral extinction and reinstatement of the CS-US association. We focused on the nucleus accumbens core (NAc) because CS-related increases in DA release have been reported for this region (Bassareo & Di Chiara, 1999; Cheng et al., 2003; Datla et al., 2002; Day et al, 2007; Robinson et al, 2001; Roitman, Stuber, Phillips, Wightman & Carelli, 2004; Stuber et al., 2008; Sunsay & Rebec, 2008). We also used FSCV after CS offset (the post-CS period) during acquisition trials to assess DA release to the US (Wise, 1978, 1982).

Methods

Subjects

Four, male Sprague-Dawley rats (370–450 g) bred from source animals supplied by Harlan Industries (Indianapolis, IN) were obtained from the departmental colony. Subjects were housed in a temperature- and humidity-controlled room on a 12-h light-dark cycle (lights on at 07:30). All animals were tested during the light cycle (10:00–16:00). All experimental procedures followed National Institutes of Health guidelines and were approved by the Indiana University Bloomington Institutional Animal Care and Use Committee (IACUC).

Apparatus

Rats were placed in a locally constructed, transparent, operant box housed inside an electrically grounded Faraday cage. The operant box (12 × 12 × 16 cm) included a hard-plastic food cup that protruded 4.5 cm into the box. Infrared photocells located inside the food cup allowed the detection of food-cup entries as the rats broke the photo beams in anticipation of the US. The food cup was located 4.5 cm above the glass-rod floor. A 10-s duration tone (1900 Hz and 70 dB) served as the CS, which was delivered through a speaker located on the wall opposite the food cup.

Surgery

In preparation for subsequent FSCV, rats were anesthetized with ketamine (80 mg/kg, i.p.) and xylazine (10 mg/kg, i.p.). After the skull was exposed, a hole was drilled 1.3 mm anterior and 1.3 mm lateral to bregma to permit access to the NAc (Paxinos & Watson, 1998). A hub, designed to mate with a locally constructed microdrive assembly on the recording day, was cemented over the hole, which was sealed with a rubber septum. A second hole, ipsilateral and 2 mm posterior to the first, was made for the reference electrode and also was equipped with a hub and rubber septum. Two animals also were equipped with a bipolar stimulating electrode (0.2-mm-diameter tips separated by 1 mm; Plastics One, Roanoke, VA, U.S.A) in the medial forebrain bundle (MFB), −4.1 anterior and +1.4 lateral to bregma and 8–9 mm ventral from the surface of the skull (Paxinos & Watson, 1998).

Pavlovian Conditioning

After at least a week of post-surgical recovery, animals were food deprived to 85% of their *ad lib* body weight.

Feeder Training and Conditioning—Each rat was first placed in the chamber and allowed to eat 0.45 mg food pellets (Noyes High Precision) delivered two at a time (0.2 s interval) into the food cup at irregular periods (10 s to 80 s) for 20 min. Rats typically received 30–40 food pellets. On subsequent days, animals received 16 CS-US pairings (trials) in each of 4–5 daily sessions with the micro-drive inserted into the hub to allow habituation to the recording set up and prevent a novelty effect on the recording day. Food pellets were delivered at the offset of the tone, permitting the animal to acquire the CS-US association. Each trial was separated by randomly intermixed intervals of four 90, four 120, four 180 and four 240 s intervals.

Extinction and Reinstatement—On the following day, the working electrode was lowered into the NAc and the animals initially received 8 to 12 CS-US pairings, separated by randomly selected inter-trial intervals (ITIs) of 90 and 120 s (four of each). When a reliable measure of performance was obtained over these trials (i.e. an elevation score representative of the previous day's performance, see below), we began extinction trials using the same ITI values. Extinction continued until performance reached zero for three consecutive trials. If extinction performance did not reach this criterion on the recording day, recording continued on the following day with a new electrode, starting with at least eight conditioning trials, followed by extinction trials and so on. After extinction performance was established, two unannounced food pellets were delivered at two occasions, separated by an interval of 120 s. In the next eight trials, the CS was presented with a 120 s ITI to determine if food deliveries reinstated the extinguished CS-US association.

Data Analysis—The number of food cup entries during the 10 s period prior to the CS (i.e., the pre-CS period) was subtracted from the number of food cup entries during the CS to calculate an elevation score, which was used as an index of learning (e.g., Bouton & Sunsay, 2003). Analysis of Variance (ANOVA) was used to determine statistical significance of the group elevation score across sessions. The analysis was considered significant if the p value was less than the 0.05 chance level.

Cohen's d index was added to each analysis to show effect size. In all cases, our d index was substantial (>0.5) owing to our within-subjects design, which reduced subject variability. The size of our d index also indicated that our sample size was sufficient in accord with IACUC requirements; additional animals were not required to show an effect.

Voltammetry

Procedure—An untreated (Thornel P-25) carbon fiber (10 μm in diameter), which served as the working electrode, was sealed in a pulled-glass capillary (150–196 μm exposed tip) according to standard procedures (Rebec et al., 1997). The electrode was first tested in citrate- phosphate buffer with 100 nM DA vs. saturated calomel reference. Cyclic voltammograms, recorded in conjunction with a two-electrode potentiostat under computer control, were obtained every 100 ms at a scan rate of 400 V s−1 (−0.4 to + 1.2). On the recording day, a freshly prepared working electrode was inserted into the microdrive for lowering into the target site for *in vivo* voltammetry. Target depth was set at 7.4 mm. A Ag/AgCl wire inserted into the second hub overlying cortex ipsilateral to the working electrode served as the reference. After the recording day, animals equipped for MFB stimulation were again anesthetized and an electrical current (60 Hz, peak-to-peak, 250 μA magnitude and 2 s duration) was applied to the MFB and a working electrode in NAc recorded voltammograms. MFB stimulation elicits transient DA release in NAc and other targets of mesolimbic DA (Garris & Rebec, 2002); voltammograms elicited by this known mechanism of DA release were used in conjunction with regression analysis to verify the DA nature of the voltammetric signal obtained during Pavlovian conditioning. We also used voltammograms obtained from in-vitro application of DA to verify the DA nature of the voltammograms.

Data Analysis—Where appropriate, data are expressed as mean ± SEM oxidation current based on background-subtracted cyclic voltammograms similar to previous work (Garris et al., 1997; Rebec et al., 1997). Each voltammogram in an ongoing data-collection file was subtracted from the average of the voltammograms collected 0.5 s earlier. Background-subtracted voltammograms obtained during conditioning sessions were subsequently compared by linear regression to peak background-subtracted voltammograms recorded during MFB stimulation and *in vitro* DA testing (Stuber, Roitman, Phillips, Carelli & Wightman, 2005). Only voltammograms having $r^2 > 0.8$ were identified as DA. Oxidation and reduction were recorded as positive and negative currents, respectively. Data analysis focused on peak increases in oxidation current that were at least two standard deviations above the values recorded over the three immediately preceding scans. Peak increases >16% above baseline were routinely identified as DA in our previous study and we used the same criterion to monitor DA signals (Sunsay & Rebec, 2008).

Histology—To mark the recording site, animals were deeply anesthetized after the last session, and current was applied to the carbon fiber to make a small lesion. This step allowed us to identify electrode placement on the last day and to confirm that more dorsal sites recorded on previous days (see above) were also confined to NAc. After transcardial perfusion with formalin, the brain was removed and subsequently frozen, sectioned, and stained with cresyl violet for histological analysis. As Figure 1 shows, electrode placements were confined to the NAc.

Results

Behavior

Acquisition—Pavlovian conditioning proceeded uneventfully. All animals learned within five acquisition sessions. A repeated measures ANOVA applied to the elevation score showed a significant effect of session, $\underline{F}(4, 12) = 8.19$, $p < 0.05$, $d = 0.75$. An identical analysis of the pre-CS scores did not show a session effect, $\underline{F}(4, 12) = 2.03$, $p = 0.14$, ruling out the possibility that the increase in elevation score reflected a pre-CS change.

Extinction and Reinstatement—Figure 2 shows behavioral performance during initial acquisition, extinction, and reinstatement trials on the recording day. Extinction was acquired as indicated by a zero elevation score for at least three successive trials. The conditioned response (CR) was reinstated following the delivery of two successive food pellets as the figure shows. A repeated-measures ANOVA, which analyzed the last extinction trial and the first three reinstatement trials, showed a significant increase in the number of food cup entries, $\underline{F}(3, 9) = 6.06$, $p < 0.05$, $d = 0.86$. An identical ANOVA performed on the pre-CS scores did not indicate a change in responding, $F < 1$, ruling out the possibility that the increase in elevation score was an arithmetical artifact.

Voltammetry

DA Verification—The bulk of DA input to NAc arises from the ventral tegmental area via the MFB. To confirm that DA release transients occur in NAc and that their identity matches the identity of the transients recorded during behavioral testing, we used a regression

analysis as in previous studies to compare the stimulation-evoked and behavior-related voltammograms (Phillips et al., 2003; Stuber et al., 2005; Sunsay & Rebec, 2008). We also included voltammograms obtained in response to 100 nM DA during *in vitro* electrode testing. Peak voltammograms from the conditioning, extinction, and reinstatement trials were found to be comparable to the voltammograms elicited by MFB stimulation and by DA *in vitro* (the smallest $r^2 = 0.91$, $p < 0.05$). Similar results (the smallest $r^2 = 0.89$, $p < 0.05$) emerged from an analysis of peak voltammograms obtained during post-CS periods. Voltammograms not resembling DA were not correlated with the voltammetric signal induced by MFB stimulation or *in vitro* testing ($r^2 = 0.40$, $p > 0.05$). Figure 3 shows an increase in oxidation current recorded from NAc upon stimulation of MFB and *in vitro* application of DA with corresponding voltammograms obtained at the time of peak current (inset); note the similarity to the voltammograms obtained *in vitro* with the same electrode after the stimulation (bottom).

The data were monitored for DA peaks during the CS periods in the last four trials of conditioning and extinction since rat performance was highest and lowest during these phases, respectively. Only the first three trials of the reinstatement phase were monitored since reinstatement performance was significantly higher than extinction performance only in these trials.

DA Signals during Acquisition of the CS—The left side of Figure 4 shows representative DA peak amplitudes from one rat during the conditioning, extinction, and reinstatement phases in the top, middle and bottom panels, respectively. Three of the four animals showed these responses. One rat failed to show significant DA peaks during conditioning trials despite behavioral evidence of learning, but DA peaks appeared during reinstatement (see below), indicating that DA was detectable in this animal.

DA Signals during Extinction and Reinstatement—Although there was a decrease in peak DA amplitudes in all rats during extinction, one rat continued to show discernible DA peaks during extinction (right side of Figure 4) despite complete elimination of the conditioned response. Free food pellets reinstated the DA peaks during the CS in all rats. Interestingly, the rat that failed to show DA peaks during acquisition trials showed DA signals during reinstatement, suggesting that habituation during conditioning may have prevented the DA response to the CS during reinforced trials at test. Non-contingent food deliveries trials may have caused dishabituation and reinstated DA activity to subsequent CS presentations.

Figure 5 shows the mean (±SEM) of the peak DA amplitude in response to the CS across acquisition, extinction, and reinstatement trials. Note the marked decrease in DA amplitude during extinction relative to the other phases. Peak height was measured as percent increase from the inflection point, which is at least two standard deviations above the mean of the preceding three successive points (Sunsay & Rebec, 2008). A repeated-measures ANOVA revealed a significant effect of phase (i.e., acquisition, extinction, reinstatement), $\underline{F}$ (2, 4) = 8.68, $d = 0.9$. Pairwise comparisons between the phases showed a significant decrease in DA peak amplitude during extinction compared to that in both acquisition and reinstatement, $\underline{F}$

$(1, 4) = 16.47$ and $\underline{F}$ $(1, 4) = 8.14$, smallest $d = 0.89$, respectively. There was no reliable difference in amplitude between acquisition and reinstatement, $\underline{F}$ $(1, 4) = 1.45$.

DA Signals after CS Offset—The upper panel in Figure 6 shows a representative pattern of DA oxidation current throughout the 10 s period of food consumption. In most trials, food consumption elicited phasic DA peaks.

Discussion

Consistent with our previous work (Sunsay & Rebec, 2008) as well as that of others (Day et al., 2007; Stuber et al., 2008), we observed DA signals during tone presentations, suggesting that phasic DA activity in the NAc serves as a neurochemical index of the associative strength of a CS. We also found that extinction reduced the amplitude of the DA signal. In fact, in some rats, extinction eliminated DA signals completely. Interestingly, DA peaks persisted in one rat despite the complete behavioral extinction of the CS-US association. This observation suggests that the neuronal representations of appetitive associations are more persistent than what behavioral measures indicate. Thus, prolonged extinction may be required to eliminate all signs of relapse.

The decrease in DA amplitude to the CS during extinction is consistent with the prediction-error model of Schultz (1998), parallels the data on the electrophysiological activity of DA neurons observed during extinction (Pan et al., 2008) and conditioned inhibition (Tobler, Dickinson & Schultz, 2003), and supports FSCV evidence based on operant conditioning and extinction procedures (Owesson-White et al., 2008). Crucially, however, the reinstatement of phasic DA signals that we observed after extinction contradicts the prediction-error model (e.g., Schultz, 1998). According to this model, DA serves as a surrogate signal for updating associative strength. After complete extinction, the associative strength of the CS-US is near zero. In other words, the prediction error model and most models of associative learning assume that extinction erases the original CS-US association. These models, therefore, do not predict anticipatory behavior or anticipatory DA signals with CS presentation at reinstatement. Simply put, presentation of a cue after its complete extinction should be no different than its first presentation in conditioning (Schultz, 1998). Thus, no DA activity would be expected at reinstatement. In contrast to this prediction, we observed that re-presentation of the reward after complete extinction of the CS-US association reinstated both behavioral anticipation and the related DA signal.

Recent prediction-error models are able to predict some of the post-extinction recovery effects. Although the tripartite extinction model of Pan et al. (2008) is able to accommodate the retrieval of CRs with the passage of time (i.e., spontaneous recovery effect), the model is unable to account for the reinstatement of DA release that we observed. The model can explain the retrieval of conditioning memory with the passage of time but not with non-contingent reward presentation. In particular, given that acquisition, extinction and reinstatement occurred within the same one-hour session, the model cannot accommodate our results because there would not be sufficient forgetting of extinction in such short amount of time.

The mathematical models that represent multidimensional physical aspects of acquisition and extinction, such as context, are better equipped to accommodate our results (Gershman et al., 2010; Redish et al., 2007). As described in the Introduction, another prediction-error model explains reinstatement in that extinction does not erase acquisition, but rather results in representing both episodes as two different categories (Redish et al, 2007). The presentation of non-contingent rewards reactivates the acquisition memory at test. The Bayesian model of learning is also successful in explaining reinstatement by a similar process that generates a new latent cause during extinction. Presentation of non-contingent rewards helps the animal infer that acquisition is active at test (Gershman et al., 2010).

The DA release that we observed following CS offset during conditioning also deserves comment. Both our previous (Sunsay & Rebec, 2008) and current work showed lingering DA activity to the presentation of rewards during conditioning, a finding that contrasts with another report (Day et al., 2007; Stuber et al., 2008). The presence of DA release at the time of expected rewards after conditioning is acquired may seem inconsistent with the prediction-error hypothesis, but as we discussed elsewhere (Sunsay & Rebec, 2008), this is a parameter-dependent prediction. When the conditioning regimen is not asymptotic, the presence of rewards should elicit DA release despite the presence of DA peaks during the CS (Pan et al., 2005). It should also be noted that the DA signals we observed after CS offset were often not time-locked to food delivery. Often, consumption of food triggered DA release, an outcome consistent with the view that DA plays some role in unconditioned reward (Wise, 1978; 1982).

Understanding the associative mechanisms of extinction and its neural correlates has important implications for many ailments such as anxiety disorders and the relapse problem in substance abuse. Mathematical models aimed at mapping DA responses to associative learning processes have renewed the interest in studies of associative learning. The exact nature of extinction nonetheless is multi-faceted and requires explanations with psychological constructs. Many mechanisms – such as memory-retrieval, emotional after effects (e.g., frustration), learning the CS-no US association, and learning an inhibitory stimulus-response association – have been suggested (Rescorla, 1997). Interestingly, the mathematical models outlined here are consistent with some of these psychological explanations. For instance, it was previously argued that omission of rewards generates frustration as an internal emotive context, which then might be associated with the physical properties of the context, thus potentially explaining the renewal effect (Bouton & Sunsay, 2001; Pearce, Redhead & Aydin, 1997). The state-partitioning process initiated by negative prediction errors in the Redish et. al. (2007) model is thus psychologically plausible. Similarly, higher-order cognitive processes such as reasoning and inference that are the hallmarks of human cognition have been argued to operate in non-human animals as well, including rats (Blaisdell, Sawa, Leising & Waldmann, 2006, but see Dwyer, Starns & Noney, 2009, for an alternative explanation).

To reach a more comprehensive understanding of DA signaling in Pavlovian extinction, alternative mechanisms of extinction and its psychologically plausible theoretical possibilities should be assessed along with mathematical models of DA responses. FSCV can play an important role in this assessment.

Acknowledgments

References

Bassareo V, Di Chiara G. Differential responsiveness of dopamine transmission to food-stimuli in nucleus accumbens shell/core compartments. Neuroscience. 1999; 89:637–641. [PubMed: 10199600]

Blaisdell AP, Sawa K, Leising KJ, Waldmann MR. Causal reasoning in Rats. Science. 2006; 311:1020–1022. [PubMed: 16484500]

Bouton ME. Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. Psychological Bulletin. 1993; 114:80–99. [PubMed: 8346330]

Bouton ME, Bolles RC. Contextual control of the extinction of conditioned fear. Learning and Motivation. 1979a; 10:445–466.

Bouton ME, Bolles RC. Role of conditioned contextual stimuli in reinstatement of extinguished fear. Journal of Experimental Psychology: Animal Behavior Processes. 1979b; 5:368–378. [PubMed: 528893]

Bouton ME, Sunsay C. Contextual control of appetitive conditioning: Influence of a contextual stimulus generated by a partial reinforcement procedure. Quarterly Journal of Experimental Psychology. 2001; 54B:109–125. [PubMed: 11393934]

Bouton ME, Sunsay C. Importance of trials versus accumulating time across trials in partially reinforced appetitive conditioning. Journal of Experimental Psychology: Animal Behavior Processes. 2003; 29:62–77. [PubMed: 12561134]

Cheng JJ, de Bruin JPC, Feenstra MGP. Dopamine efflux in nucleus accumbens shell and core in response to appetitive classical conditioning. European Journal of Neuroscience. 2003; 18:1306–1314. [PubMed: 12956729]

Datla KP, Ahier RG, Young AMJ, Gray JA, Joseph MH. Conditioned appetitive stimulus increases extracellular dopamine in the nucleus accumbens of the rat. European Journal of Neuroscience. 2002; 16:1987–1993. [PubMed: 12453062]

Day JJ, Stuber MF, Wightman RM, Carelli RM. Associative learning mediates dynamic shifts in dopamine signaling within the nucleus accumbens. Nature Neuroscience. 2007; 10:1020–28.

Dwyer DM, Starns J, Honey RC. Causal reasoning” in rats: A reappraisal. Journal of Experimental Psychology: Animal Behavior Processes. 2009; 35:578–586. [PubMed: 19839709]

Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CAA, Clinton SM, Philips PEM, Akil H. A selective role for dopamine in reward learning. Nature. 2011; 469:53–57. [PubMed: 21150898]

Garris PA, Rebec GV. Modeling fast dopamine neurotransmission in the nucleus accumbens during behavior. Behavioral Brain Research. 2002; 137:47–63.

Gershman SJ, Blei DM, Niv Y. Context, learning and extinction. Psychological Review. 2010; 117:197–209. [PubMed: 20063968]

Hart AS, Rutledge RB, Glimcher PW, Phillips PEM. Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. The Journal of Neuroscience. 2014; 34:698–704. [PubMed: 24431428]

Ikemoto S, Panksepp J. The role of nucleus accumbens dopamine in motivated behavior: A unifying interpretation with special reference to reward-seeking. Brain Research Reviews. 1999; 31:6–41. [PubMed: 10611493]

Mirenowicz J, Schultz W. Importance of unpredictability for reward responses in primate dopamine neurons. Journal of Neurophysiology. 1994; 72:1024–1027. [PubMed: 7983508]

Mirenowicz J, Schultz W. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. Nature. 1996; 379:449–451. [PubMed: 8559249]

Owesson-White CA, Cheer JF, Beyene M, Carelli RM, Wightman RM. Dynamic changes in accumbens dopamine linked to learning of intra-cranial self-stimulation. Proceedings of the National Academy of Sciences. 2008; 105:11957–62.

Pan WX, Schmidt R, Wickens JR, Hyland BI. Dopamine cells respond to predicted events during classical conditioning: Evidence for eligibility traces in the reward-learning network. Journal of Neuroscience. 2005; 25(26):6235–6242. [PubMed: 15987953]

Pan WX, Schmidt R, Wickens JR, Hyland B. Tripartite mechanism of Extinction suggested by dopamine neuron activity and temporal difference model. The Journal of Neuroscience. 2008; 28:9619–9631. [PubMed: 18815248]

Paxinos, G.; Watson, C. The rat brain in stereotaxic coordinates. 4. Academic Press; San Diego: 1998.

Pearce JM, Redhead ES, Aydin A. Partial reinforcement in appetitive conditioning with rats. Quarterly Journal of Experimental Psychology. 1997; 50B:273–294. [PubMed: 9421975]

Phillips PEM, Stuber GD, Heien MLAV, Wightman RM, Carelli RM. Subsecond dopamine release promotes cocaine seeking. Nature. 2003; 422:614–618. [PubMed: 12687000]

Rebec GV, Christensen JRC, Guerra C, Bardo MT. Regional and temporal differences in real-time dopamine efflux in the nucleus accumbens during free-choice novelty. Brain Research. 1997; 776:61–67. [PubMed: 9439796]

Redgrave P, Gurney K. The short-latency dopamine signal in discovering novel actions. Nature Reviews Neuroscience. 2006; 7:967–974.

Redish AD, Jensen S, Johnson A, Kurth-Nelson Z. Reconciling reinforcement learning models with behavioral extinction and renewal: Implications for addiction, relapse and problem gambling. Psychological Review. 2007; 114:784–805. [PubMed: 17638506]

Rescorla RA. Response-inhibition in extinction. Quarterly Journal of Experimental Psychology. 1997; 50B:238–52.

Rescorla RA. Spontaneous recovery. Learning and Memory. 2004; 11:501–509. [PubMed: 15466300]

Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black, AH.; Prokasy, WF., editors. Classical conditioning II: Current research and theory. New York: Appleton-Century-Crofts; 1972. p. 64-99.

Robinson DL, Phillips PEM, Budygin EA, Traftin BJ, Garris PA, Wightman RM. Sub-second changes in accumbal dopamine during sexual behavior in male rats. NeuroReport. 2001; 12:2549–2552. [PubMed: 11496146]

Roitman MF, Stuber GD, Phillips PEM, Wightman RM, Carelli RM. Dopamine operates as a subsecond modulator of food seeking. Journal of Neuroscience. 2004; 24:1265–1271. [PubMed: 14960596]

Schultz W. Predictive reward signal of dopamine neurons. Journal of Neurophysiology. 1998; 80:1–27. [PubMed: 9658025]

Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997; 275:1593–1599. [PubMed: 9054347]

Stuber GD, Klanker M, de Ridder B, Bowers MS, Joosten RN, Feenstra MG, Bonci A. Reward-predictive cues enhance excitatory synaptic strength onto midbrain dopamine neurons. Science. 2008; 321:1690–1692. [PubMed: 18802002]

Stuber GD, Roitman MF, Phillips PEM, Carelli RM, Wightman RM. Rapid dopamine signaling in the nucleus accumbens during contingent and noncontingent cocaine administration. Neuropsychopharmacology. 2005; 30:853–863. [PubMed: 15549053]

Stuber GD, Wightman RM, Carelli RM. Extinction of cocaine self-administration reveals functionally and temporally distinct dopaminergic signals in the nucleus accumbens. Neuron. 2005; 46:661–669. [PubMed: 15944133]

Sunsay C, Rebec GV. Real-time dopamine efflux in the nuclues accumbens core during Pavlovian conditioning. Behavioral Neuroscience. 2008; 122:358–367. [PubMed: 18410174]

Tobler PN, Dickinson A, Schultz W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. The Journal of Neuroscience. 2003; 23:10402–10410. [PubMed: 14614099]

Wise RA. Catecholamine theories of reward: A critical review. Brain Research. 1978; 152:215–247. [PubMed: 354753]

Wise RA. Neuroleptics and operant behavior: The anhedonia hypothesis. Behavioral Brain Sciences. 1982; 5:39–87.

Fig. 1.
Schematic depiction of electrode placements compiled for display in a coronal section of
NAc shown at 1.60, 1.20, 1.00 and 0.70 mm anterior to bregma according to Paxinos and
Watson (1998). Filled and open circles represent electrode placements for Experiments 1
and 2, respectively. Placements in all tested animals were marked on the last day of
recording and subsequently verified by histological analysis. From *The Rat Brain in
Stereotaxic Coordinates* (4th ed.), by G. Paxinos and C. Watson, 1998, San Diego, CA:
Academic Press. Adapted with permission.
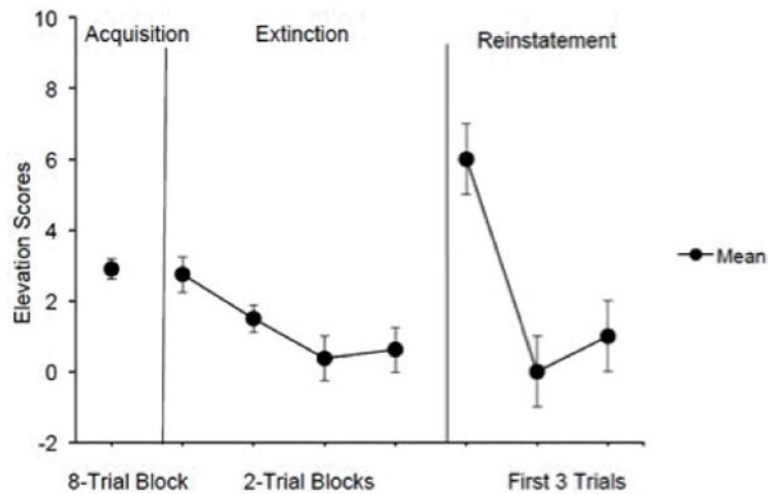
Fig. 2.
Mean elevation scores calculated as the number of food cup entries in the pre-CS period subtracted from the number of food cup entries in the CS period. Acquisition elevation scores were averaged across all sessions. Extinction performance is shown as elevation scores averaged over two-trial blocks. Only the first-day extinction performance was calculated into the group data for the figure. Non-zero performance is shown because one rat reached extinction criterion on the second recording day. This rat had higher extinction performance than the others on the first day, which made the average over the extinction blocks larger than zero. Reinstatement performance is shown across three trials. Brackets indicate SEM.
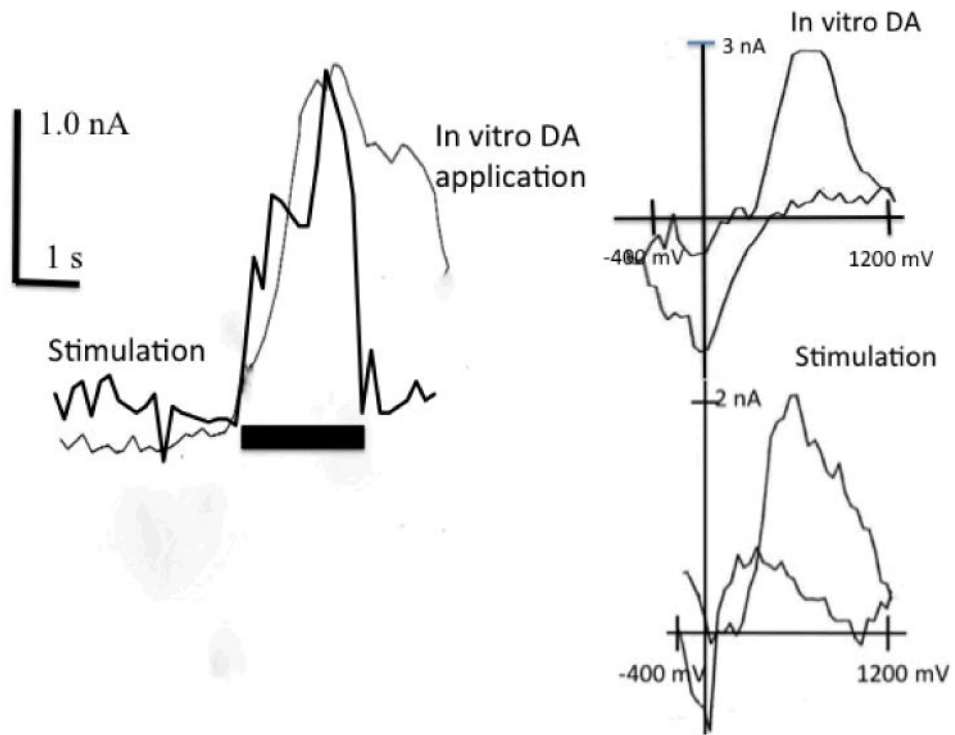
Fig. 3.
Representative examples of transient DA oxidation signals elicited by MFB stimulation and in vitro DA application. The horizontal bar indicates the onset of 2 s bath DA application and stimulation. Scale bars indicate 1 nA (vertical) and 1 s (horizontal). Inset: background subtracted cyclic voltammograms obtained at the peak of the response.
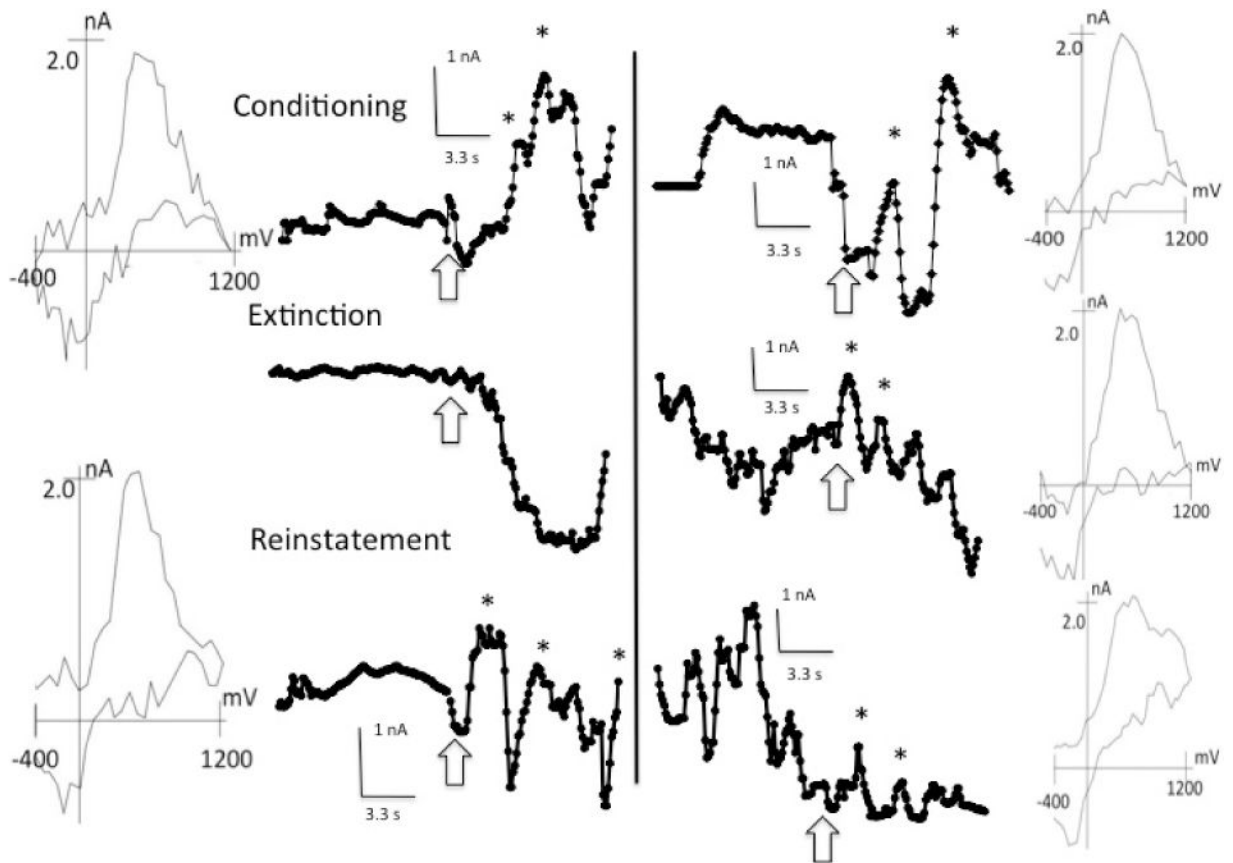
Fig. 4.
DA transients recorded during the 10 s CS and 10 s pre-CS periods during the acquisition, extinction, and reinstatement trials for Rats 4 and 2 in the left and right columns, respectively. DA transients, indicated by asterisks, were confirmed by corresponding background subtracted voltammograms shown in the left-most and right-most columns for Rats 4 and 2, respectively. Arrows indicate the onset of the 10 s tone period. Only first-peak voltammograms are presented.
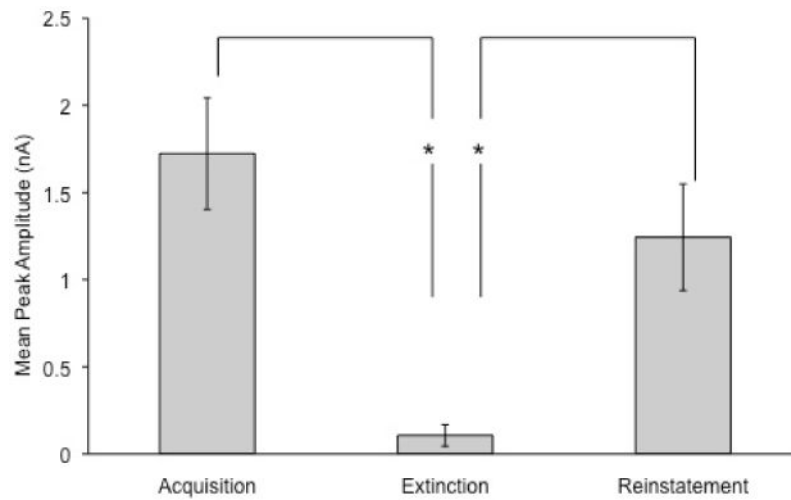
Fig. 5.
Mean (±SEM) of DA peak amplitudes to the CS across acquisition, extinction, and reinstatement trials. Conditioning produced DA transients in most animals and extinction reduced the amplitudes of the DA transients that reappeared during reinstatement. Analysis was limited to transients 16% above background and voltammetrically identified as DA. Brackets indicate SEM. Asterisks indicate significant differences between phases.
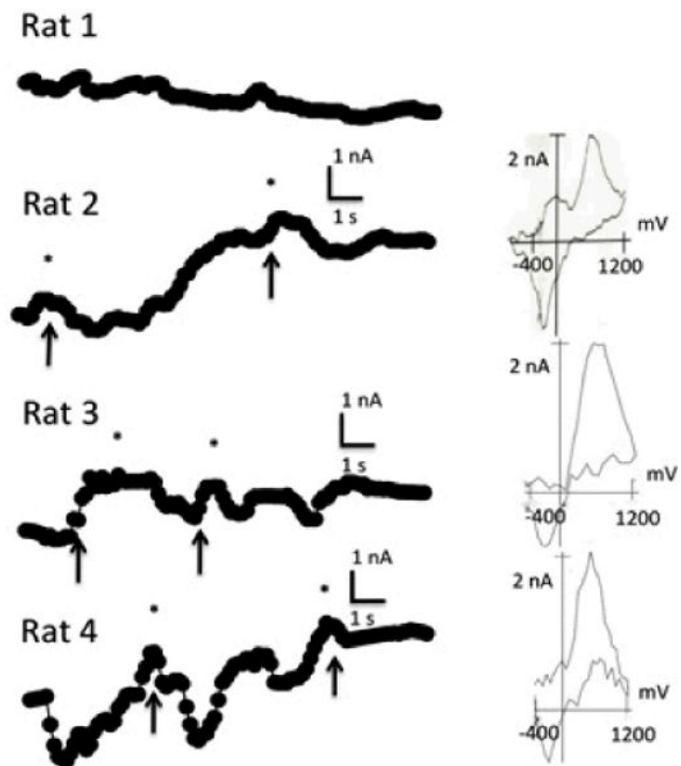
Fig. 6.
Voltammetric signals recorded for all rats during the 10 s US period associated with the acquisition trials. Arrows indicate food cup entries. DA transients, indicated by asterisks, were confirmed by corresponding background subtracted voltammograms shown in the right-most column. Note the lack of transients in Rat 1.