

# Phasic Mesolimbic Dopamine Signaling Precedes and Predicts Performance of a Self-Initiated Action Sequence Task

Kate M. Wassum, Sean B. Ostlund, and Nigel T. Maidment

**Background:** Sequential reward-seeking actions are readily learned despite the temporal gap between the earliest (distal) action in the sequence and the reward delivery. Fast dopamine signaling is hypothesized to mediate this form of learning by reporting errors in reward prediction. However, such a role for dopamine release in voluntarily initiated action sequences remains to be demonstrated.

**Methods:** Using fast-scan cyclic voltammetry, we monitored phasic mesolimbic dopamine release, in real time, as rats performed a self-initiated sequence of lever presses to earn sucrose rewards. Before testing, rats received either 0 ( $n = 11$ ), 5 ( $n = 11$ ), or 10 ( $n = 8$ ) days of action sequence training.

**Results:** For rats acquiring the action sequence task at test, dopamine release was strongly elicited by response-contingent (but unexpected) rewards. With learning, a significant elevation in dopamine release preceded performance of the proximal action and subsequently came to precede the distal action. This predistal dopamine release response was also observed in rats previously trained on the action sequence task, and the amplitude of this signal predicted the latency with which rats completed the action sequence. Importantly, the dopamine response to contingent reward delivery was not observed in rats given extensive pretraining. Pharmacological analysis confirmed that task performance was dopamine-dependent.

**Conclusions:** These data suggest that phasic mesolimbic dopamine release mediates the influence that rewards exert over the performance of self-paced, sequentially-organized behavior and sheds light on how dopamine signaling abnormalities may contribute to disorders of behavioral control.

**Key Words:** Dopamine, flupenthixol, free-operant, reinforcement learning, reward, ventral striatum

A long-standing challenge in behavioral science, termed the distal reward (1) or credit assignment problem (2), has been to explain how action sequences are learned, given that one initiates a sequence by performing an action that never directly earns reward. There is growing evidence that dopamine is involved in mediating this influence of rewards over distal events. It is imperative to understand the role of dopamine in action sequence learning and in the control of self-initiated behavior, more generally, since these actions appear to be particularly disrupted by disorders involving alterations in striatal dopaminergic transmission, such as Parkinson's disease (3) and addiction (4–6).

Studies recording dopamine cell activity have shown that reward presentation elicits a firing burst that shifts to reward-predictive cues (7–9), or discriminative stimuli (10,11), with training (12). Fast-scan cyclic voltammetry studies have found that phasic mesolimbic dopamine release also backpropagates from reward to reward-predictive cues during Pavlovian learning (13,14) or to a discriminative stimulus during discrete-trial instrumental tasks (15,16). In paradigms involving a chain of reward-predictive cues, dopamine cell activity also shifts from proximal to distal cues (17,18). Such findings support the hypothesis that dopamine re-

ports reward prediction errors, i.e., discrepancies between the observed and expected values of rewards and cues (19,20). This signal is considered important for the acquisition of complex reward-seeking behaviors (21,22). Interestingly, dopamine receptor blockade has recently been shown to reduce response likelihood during a self-initiated reward-seeking action sequence (23), and accumbal dopamine depletions disrupt action performance, when many actions are required to obtain reward (24), which is consistent with a large literature implicating dopamine in incentive motivation (25–28), the process that allows reward-predictive cues to invigorate reward-seeking actions. Such findings raise the possibility that task-related dopamine signaling also exerts a motivational influence over reward-seeking actions.

Although the simplifying assumptions common to reinforcement learning theory make it difficult to generate predictions about the characteristics of reward prediction error (and phasic dopamine) signaling in self-paced, free-operant instrumental tasks, there are limited data to suggest that in such situations a similar pattern of dopamine responses emerges with learning. Indeed, both dopamine neuron firing (29,30) and phasic mesolimbic dopamine release (31,32) precede well-learned self-initiated actions, consistent with a backpropagating reward prediction error signal. However, studies of dopamine release during free-operant learning have typically used cocaine reward (31,32), which may not generate normal prediction error signals (33), given that it elicits dopamine when predicted (31) and over time undoubtedly alters dopaminergic function (34). Whether phasic mesolimbic dopamine release shows the properties of a prediction error signal under naturalistic free-operant conditions therefore remains unanswered. Moreover, given the prominent role that modern reinforcement learning models attribute to phasic dopamine in action sequence learning and a reward's ability to influence distal events, it is surprising how little is known about the characteristics of dopamine release in such situations. Uncovering this informa-

From the Department of Psychology (KMW); Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior (KMW, SBO, NTM); and Brain Research Institute (KMW, SBO, NTM), University of California Los Angeles, Los Angeles, California.

Authors KMW and SBO contributed equally to this work.

Address correspondence to Kate Wassum, Ph.D., University of California Los Angeles, Department of Psychology, Los Angeles, CA 90095; E-mail: kwassum@ucla.edu.

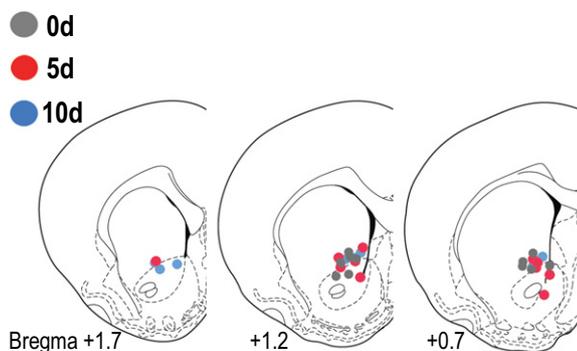
Received Oct 18, 2011; revised Nov 19, 2011; accepted Dec 20, 2011.

tion is vital if modern reinforcement learning concepts are to be used to explain the neural mechanisms of normal and aberrant decision-making.

We used fast-scan cyclic voltammetry to provide an initial characterization of phasic dopamine release in the ventral striatum/nucleus accumbens of rats performing an unsignaled, self-paced multi-action sequence task for sucrose reward. Given the combined findings that phasic dopamine activity backpropagates from reward through a series of Pavlovian cues (17,18) and can come to precede first-order self-initiated instrumental actions for cocaine reward (31,32,35), we assessed whether phasic dopamine release shifted with learning from the reward itself to increasingly more distal events. Given the large body of evidence implicating mesolimbic dopamine in incentive motivation (25,36,37), we also tested the relationship between task-related dopamine responses and action sequence performance.

## Methods and Materials

Male Sprague Dawley rats ( $n = 46$ ; Charles River Laboratories, Wilmington, Massachusetts) served as the subjects for these experiments. For the fast-scan cyclic voltammetry experiments, rats were trained on a sequence of lever-pressing actions to earn sucrose pellet rewards (Bioserv, Frenchtown, New Jersey). Briefly, the behavioral paradigm (Supplement 1) required rats to perform a fixed sequence of two different lever press actions to earn sucrose pellets, such that one action was temporally distal and the other temporally proximal to reward delivery. The distal lever was continuously available and, when pressed, resulted in the insertion of the proximal lever into the chamber. Pressing the proximal lever resulted in the delivery of a pellet and caused that lever to be retracted. Importantly, this task was self-paced, in that the subject could control both the initiation of each sequence and the speed with which the sequence of actions was performed (i.e., trial latency). Before testing, rats received either 0 ( $n = 11$ ), 5 ( $n = 11$ ), or 10 ( $n = 8$ ) days of action sequence training. During the test, phasic dopamine concentration changes in the ventral portion of the neostriatum/nucleus accumbens, a region previously implicated in reward-motivated behavior (38–40), were monitored with fast-scan cyclic voltammetry (Supplement 1) while rats earned a total of 30 sucrose pellets rewards through their action sequence performance. All recording sites were verified with histological procedures (Figure 1).



**Figure 1.** Schematic representation of recording sites. Numbers to the lower right of each section represent the anterior-posterior distance (mm) from bregma of the section. Gray circles represent electrode placements from the 0-day training group, red circles show placements for rats in the 5-day group, and blue circles show placements for rats in the 10-day training group. (Line drawings of coronal section are reproduced from Paxinos and Watson [59], with permission from Elsevier, copyright 1998).

In a separate set of rats, we examined the effects of flupenthixol (Sigma Aldrich, St. Louis, Missouri) on sequence performance after either 0, 5, or 10 days of action sequence training. Flupenthixol (5 mg/kg/mL intraperitoneal) or saline control was administered 1 hour before a test in which rats were allowed to respond on the action sequence to earn up to 30 sucrose pellet rewards.

For full methodological details, see Supplement 1.

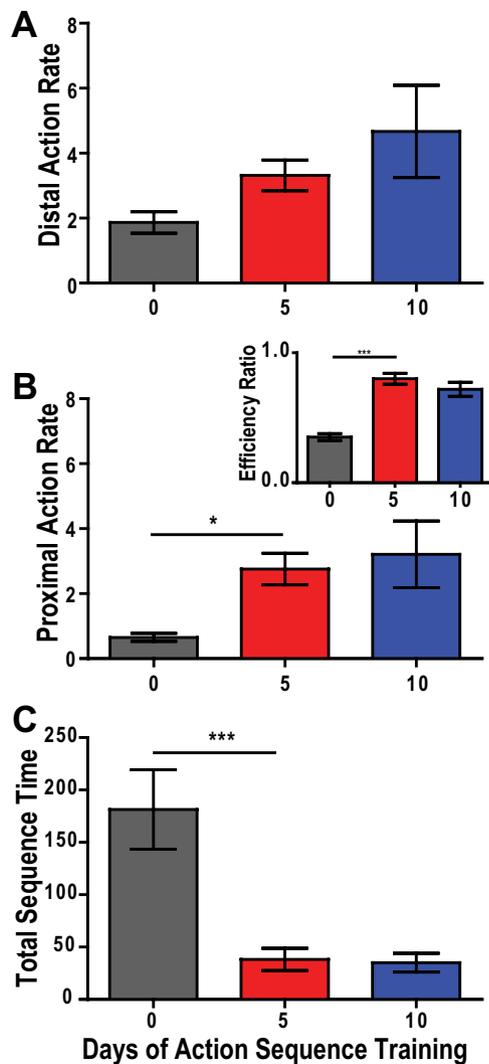
## Results

Using several measures of task performance (distal and proximal action rate, task efficiency, and the total time to complete the sequence), we found that rats given 5 or 10 days of training displayed similar levels of performance and that task performance was significantly better in these groups relative to rats that did not receive any pretest sequence training. Analysis of rats' distal lever press rate (Figure 2A) revealed a main effect of training [ $F(2,27) = 3.32, p = .05$ ], suggesting that distal action rate was significantly higher in groups with more training. Similarly, there was a significant increase in proximal action rate across training groups [ $F(2,27) = 5.88, p = .007$ ; Figure 2B]. Importantly, the proximal/distal action ratio, a measure of task efficiency executed, also showed a significant improvement across training groups [ $F(2,27) = 38.24, p < .0001$ ; Figure 2B, inset]. Dunnett's post hoc analyses of these data showed that task efficiency was significantly better in the 5-day training group ( $p < .001$ ) relative to the 0-day group and that no further improvement occurred after 10 days of sequence training (5-day vs. 10-day,  $p > .05$ ). Not only did training result in an increase in task efficiency, there was also an effect of training on the average time it took rats to complete each sequence [ $F(2,27) = 11.23, p = .0003$ ; Figure 2C]. Post hoc analyses revealed that the average action sequence time was significantly longer in rats without sequence pre-training relative to rats trained for 5 days on the sequence before test ( $p < .001$ ), while rats trained for 10 days did not perform differently than those trained for 5 days ( $p > .05$ ).

### Within a Single Learning Session, Phasic Mesolimbic Dopamine Release Comes to Precede First the Proximal Then the Distal Action in a Lever-Pressing Sequence Earning Reward

As mentioned above, mesolimbic dopamine release is hypothesized to backpropagate from reward to more distal elements in a sequence with learning. Our results appear to be consistent with this idea (see Figure 3 for representative data). Figure 4 shows the dopamine concentration change averaged across rats in the 5 seconds before and after the distal (Figure 4A) and proximal (Figure 4B) lever press at the beginning (representative example shown in Figure 3, left), middle, and end (representative example Figure 3, second panel) of the acquisition session for 0-day training group rats. Early in training, when reward delivery was relatively unexpected, mesolimbic dopamine release peaked following reward delivery (which occurred immediately following the proximal lever press, marked "Reward" in Figure 4B). Although rats had no training on the full action sequence before test, they had received brief training on the proximal lever-reward contingency, and as such, the proximal lever insertion likely served as a reward-predictive cue and elicited a small dopamine response. In the middle of the acquisition session, the maximal dopamine signal amplitude occurred earlier, coincident with the proximal press. By the end of the session, when the rats had presumably come to expect response-contingent reward delivery, there was a dopamine increase before the distal action.

Peak dopamine concentration change to each event in the sequence (preceding distal action, to the lever insertion cue, preceding proximal action, and following reward delivery) averaged



**Figure 2.** Action sequence performance during the fast-scan cyclic voltammetry tests. Rats were given either 0 (gray), 5 (red), or 10 (blue) days of action sequence training before the fast-scan cyclic voltammetry test. **(A)** The average rate of lever pressing on the distal lever. **(B)** The average rate of lever pressing on the proximal lever at test. The inset shows the efficiency ratio of the total number of proximal presses/the total number of distal presses. An efficiency ratio of 1 would indicate optimal performance. **(C)** The average time taken to complete a sequence (totaling the time from initial distal press to the next initiating distal press, as described in the text above). Error bars indicate  $\pm 1$  standard error of the mean. \* $p < .05$ , \*\*\* $p < .001$ .

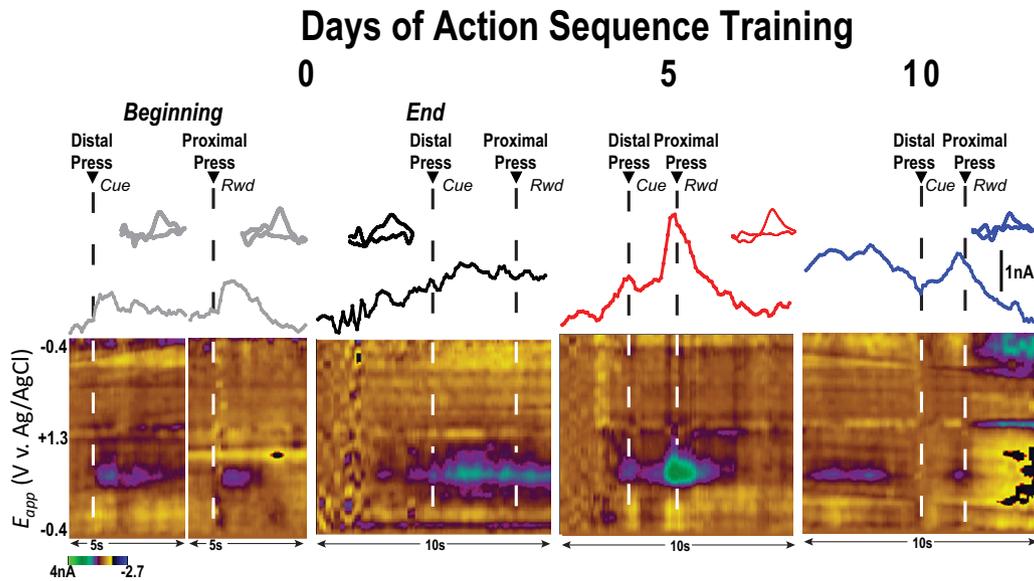
across session phase (either the first five [beginning], middle five [middle], or last five [end] trials of the acquisition session) and then averaged across subjects are presented in Figure 4C. Analysis of these data revealed no significant effect of event [ $F(3,30) = 1.30$ ,  $p = .29$ ], but did show a significant session phase effect [ $F(2,20) = 8.43$ ,  $p = .002$ ], as well as a significant phase by event interaction [ $F(6,60) = 13.08$ ,  $p = .01$ ], demonstrating that the dopamine release in response to each task event was differentially altered by action sequence learning. Indeed, individual analysis of the dopamine signal associated with each event in the beginning of the session confirmed a significant effect of event [ $F(3,30) = 5.18$ ,  $p = .005$ ], with post hoc analysis revealing that the dopamine response to the reward delivery was significantly higher than to that preceding the distal action ( $p < .01$ ), to the lever cue ( $p < .05$ ), and preceding

the proximal action ( $p < .01$ ). This was not the case in the middle and end of the session; there was only a marginally insignificant effect of event on dopamine peak concentration in the middle of the session [ $F(3,30) = 2.6$ ,  $p = .07$ ] and no effect of event at the end of the acquisition session [ $F(3,30) = .66$ ,  $p = .58$ ]. Importantly, analysis of the dopamine response to each event individually in the beginning, middle, and end of the acquisition session supports the notion that task-related dopamine signaling was modulated by training. There was a significant effect of time on the peak dopamine concentration change before the proximal response [ $F(3,30) = 6.53$ ,  $p = .007$ ], with post hoc analysis confirming that this response was significantly increased in both the middle ( $p < .05$ ) and end ( $p < .01$ ) of the session relative to the beginning of the session. Similarly, there was a significant effect of time on the predistal dopamine response [ $F(3,30) = 4.62$ ,  $p = .02$ ]; however, post hoc analyses here showed that, relative to the beginning, the amplitude of this response was increased only at the end of the session ( $p < .05$ ). Thus, it appears that for rats learning to perform a new sequence of actions, the mesolimbic dopamine response increased first during the period immediately before proximal action performance and then during the period just before distal action performance, consistent with a backpropagating, reward-prediction error profile. Importantly, rather than dopamine being solely elicited by overt cues or events, it also came to precede the rats' initiation of the action sequence, which is notable for a self-paced task in which rats' reward seeking was voluntary.

#### Across Training Groups, There Is a Shift in the Phasic Mesolimbic Dopamine Release Pattern Indicating Backpropagation from Reward to More Distal Action Sequence Elements

This shift in dopamine to more distal elements, noticed within the group acquiring the action sequence task, appeared to be followed by longer-term changes in task-related dopamine signaling, which were apparent when assessed across groups of rats with differing sequence training levels. Figures 5A and 5B show the dopamine concentration change in the 5 seconds before and after each distal (Figure 5A) and proximal (Figure 5B) lever press averaged across the 30-trial session for each rat and then averaged across rats for subjects in the 0-day, 5-day, or 10-day training groups. As is clear from this figure and the representative examples shown in Figure 3, mesolimbic dopamine was elevated both before and after the distal and proximal actions in all three groups. However, the amplitude and pattern of these dopamine concentration changes differed across training groups; the phasic dopamine signal was more prominent in the end stages of the sequence (proximal lever press and following reward delivery) in rats acquiring the sequence at test (0-day group) and became preferentially associated with more distal elements of the sequence in extensively trained rats (10-day group).

Statistical analysis of the peak dopamine release associated with each event in the sequence (Figure 5C) further supports the notion that dopamine levels surrounding the major events within the sequence critically depended on the extent of training before testing [event  $\times$  training group interaction:  $F(6,81) = 46.42$ ,  $p = .005$ ]. Dopamine levels were significantly greater during the reward delivery compared with the predistal press period ( $p < .05$ ) for rats without sequence pretraining, whereas no such difference was observed in rats given 5 days of training before testing ( $p > .05$ ). Rats given extensive pretraining (10-day group) showed the opposite effect, exhibiting a larger dopamine response before performing the distal action than to reward delivery ( $p < 0.05$ ). There were no



**Figure 3.** Examples of phasic mesolimbic dopamine release during action sequence learning and performance. Mesolimbic phasic dopamine release was recorded with fast-scan cyclic voltammetry in rats learning to perform a two-action sequence to earn sucrose rewards after either 0, 5, or 10 days of pretraining on the action sequence. Current at the peak oxidative potential of dopamine is plotted as a function of time. The insets show the background-subtracted cyclic voltammogram from the peak of current for each trace. Below are two-dimensional color plots of cyclic voltammograms over time. The y axis indicates the applied potential, the x axis shows time in seconds, and current is plotted in false color (see lower left scale). Dashed lines mark the time of either the distal or proximal press, as denoted above the current trace. Under the press demarcation, Cue refers to the insertion of the proximal lever following the distal lever press and Rwd refers to the delivery of the sucrose pellet reward after the proximal lever press. For the example, from the groups of rats acquiring the task for the first time at test (0-day group), the representative example is taken from the first action sequence trial and the last action sequence trial for the same subject. The first action sequence trial is broken into separate 5-second bins due to the fact that the rat's latency to press the proximal lever after the distal press was greater than 5 seconds. Ag/AgCl, silver/silver chloride; Rwd, reward.

main effects of either event [ $F(3,81) = 15.02, p = .36$ ] or training group [ $F(2,81) = 1.29, p = .29$ ].

Rats were intermittently given noncontingent sucrose pellets before each test to compare dopamine responses to these unexpected rewards with those earned by lever pressing. This analysis (Figure 6) revealed a main effect of expectancy [ $F(1,27) = 10.83, p = .002$ ], suggesting that the dopamine response to the earned, and therefore expected, reward was less than to unexpected reward delivery. Although there was no training group effect [ $F(2,27) = 1.56, p = .23$ ] or interaction between training group and expectancy [ $F(2,27) = .60, p = .55$ ], Bonferroni post hoc analysis revealed that the effect of expectancy was significant only in the 10-day training group ( $p < .05$ ), i.e., only in the most well-trained rats did earned reward elicit significantly less dopamine release than unexpected reward delivery. Importantly, no group differences were observed in the magnitude of the dopamine response to unexpected reward [ $F(2,29) = .42, p = .66$ ], demonstrating that the group differences described above were specific to predictable rewards and depended on rats' training history rather than other potential differences across groups (e.g., electrode sensitivity).

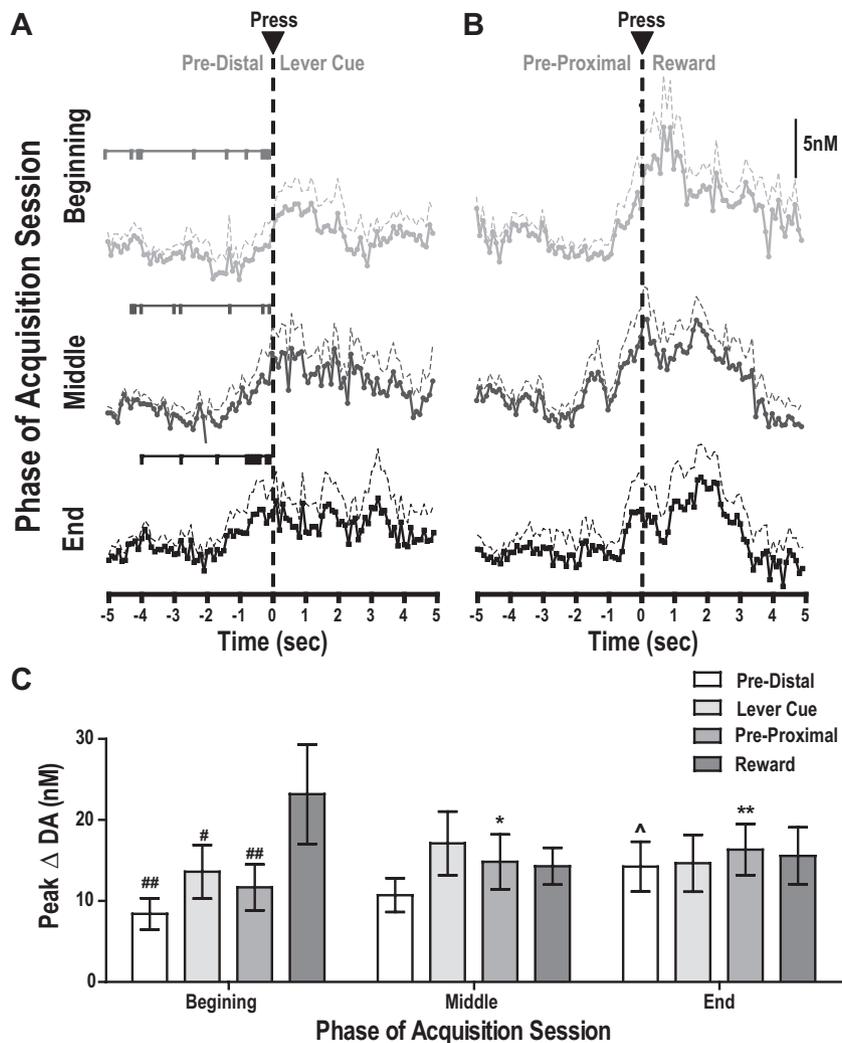
Taken together, these data show that the rapid, within-session modulation of task-related dopamine signaling observed in rats learning to perform the action sequence continues to occur over sessions, as animals become proficient in the task. Moreover, there appears to be an interim learning phase in which phasic dopamine is elevated to each event in the sequence (5-day group) before transitioning from the reward to more distal elements (10-day group).

Importantly, as can be seen in Figure 5A, the dopamine peak preceding the distal action is tightly time-locked to the actual distal lever press in both the 0- and 5-day training groups (see tick line Figure 5A, top and middle). In both these groups, dopamine levels

peaked between 3.8 and 0 seconds before the distal lever press, and for most rats, the peak occurred within 1 second before the press. The average time between these dopamine peaks and the distal lever press was not significantly different between the 0-day and 5-day training groups [ $t(20) = .31, p = .76$ ]. In the group receiving extended training (10 days) on the action sequence, time locking of the dopamine peak preceding the distal action was more variable across rats (ranging from 4.9 to .3 seconds before the press; see tick line Figure 5A, bottom) and occurred much earlier, on average, than it did for the 5-day group that had less training but showed comparable behavioral performance [ $t(17) = 2.17, p = .04$ ]. This pattern explains the apparent slow rise in average dopamine levels before the distal action in the 10-day group (Figure 5A). Rather than reflecting a consistent pattern across subjects, the averaged data obscure individual differences in preresponse dopamine signaling. Thus, it appears that the temporal relationship between phasic dopamine transients and initiation of sequence performance became decoupled for rats receiving extensive training.

### Phasic Mesolimbic Dopamine Predicts Task Performance

Our finding that mesolimbic dopamine release actually came to precede the rats' initiation of the action sequence suggests that dopamine release may not simply be a response to overt task cues, but may also reflect a motivational component of task performance. Indeed, not only was there an evolution of phasic mesolimbic dopamine release before action sequence performance (i.e., before the distal press), the magnitude of this effect predicted task performance. Figure 7A presents the concentration of the dopamine peak before the distal action for each of the 30 total events (shaded to reflect 5-trial bins) relative to the average amount of time it took to complete the sequence, averaged across rats in the 0-day training group. The magnitude of the dopamine peak pre-



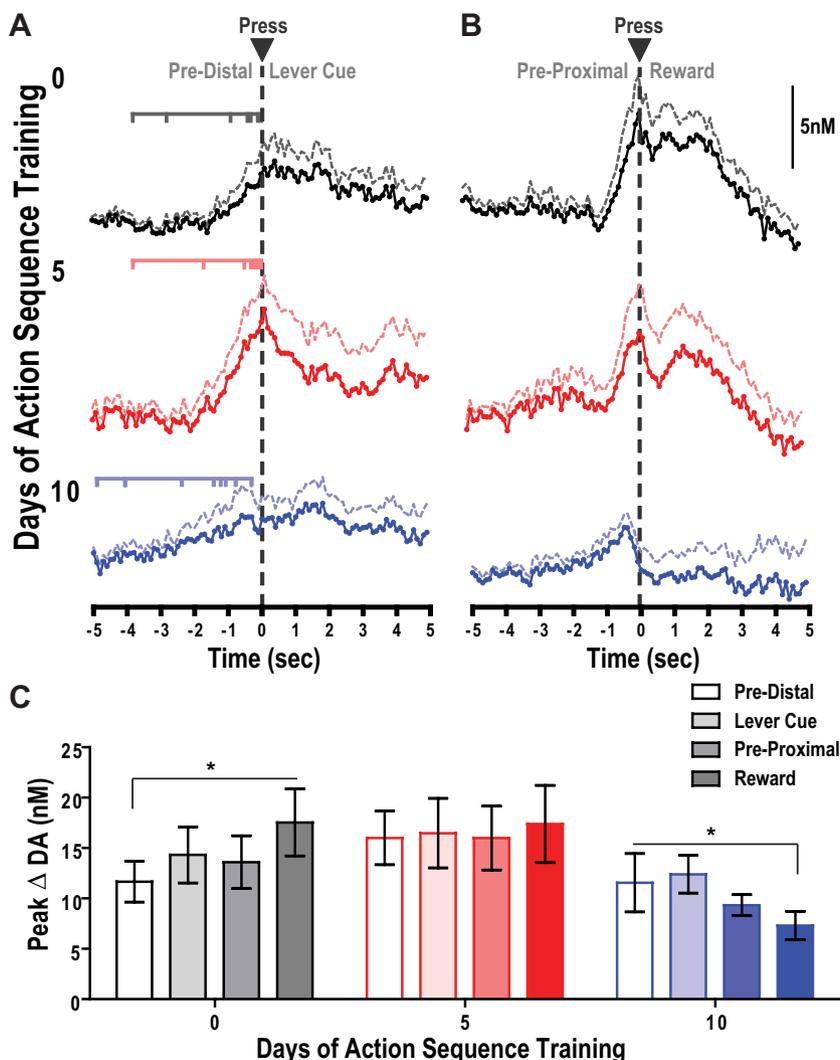
**Figure 4.** Phasic mesolimbic dopamine release comes to precede the proximal action then to precede the distal lever press during initial action sequence learning. **(A, B)** Change in dopamine concentration in the 5 seconds before and after the distal **(A)** or proximal **(B)** lever press was averaged across subjects in the 0-day group of rats acquiring the action sequence task for the first time at test. The lightest gray trace (top) represents the dopamine concentration change during the beginning of the session, averaged across the first five action sequence trials; the darker gray trace (middle) represents the dopamine concentration change averaged across the middle five action sequence trials; and the black trace (bottom) represents the same for the end of the session, the last five trials. The scale bar represents a 5 nmol/L dopamine concentration change. Tick marks in **(A)** represent the average time, for the first five (light gray, top), middle five (gray, middle), or last five (black, bottom) trials at which the dopamine concentration change peaked before the distal press. **(C)** The maximal dopamine peak concentration change preceding the distal action, upon the lever insertion cue, preceding the proximal action, and following reward delivery averaged across the first five (beginning), middle five (middle), or last five (end) trials for each rat and then averaged across rats in the 0-day training group only. Asterisks and carets mark changes in peak dopamine concentration at the middle and end of the sequence acquisition session relative to the beginning of the session. Number signs demarcate a significant difference in peak dopamine concentration preceding the distal or proximal actions or to the lever cue within each phase (i.e., beginning) relative to the reward-related dopamine response. 0-day training group,  $n = 11$ . Dashed lines and error bars indicate  $\pm 1$  standard error of the mean. \* $p < .05$ , \*\* $p < .01$ ; ^ $p < .05$ ; # $p < .05$ ; ## $p < .05$ . DA, dopamine.

ceding each distal action was significantly negatively correlated with action sequence time [ $R(30) = -.49, p = .006$ ]. A similar relationship was also apparent between subjects across all three training groups (Figure 7B). Statistical analysis of these data, controlling for training group, also revealed a significant negative correlation [ $R(27) = -.44, p = .02$ ]. Thus, it appeared that the more dopamine released before initiation of the action sequence, the quicker the rat completed the sequence. This finding suggests that phasic dopamine is not solely serving as a prediction error signal during action sequence learning but may also be related to the influence of incentive motivation on task performance. Interestingly, there were no significant correlations between the dopamine release amplitude to any of the action sequence elements and the rats' response

rate on either lever (Table S1 in Supplement 1), indicating that the rats' overt behavioral output level was not associated with phasic mesolimbic dopamine activity and was, in this sense, dissociable from their action sequence performance.

#### Action-Sequence Learning and Performance Are Attenuated Under Dopamine Receptor Blockade

To further explore the relationship between dopamine signaling and self-initiated action sequence performance, we conducted an experiment in a separate group of animals to assess the sensitivity of the current task to dopamine receptor blockade. Naive rats were given 10 days of training on the action sequence. Rats were pretreated with the nonspecific dopamine receptor antagonist flupenthixol (.5 mg/kg intraperitoneal) or vehicle be-



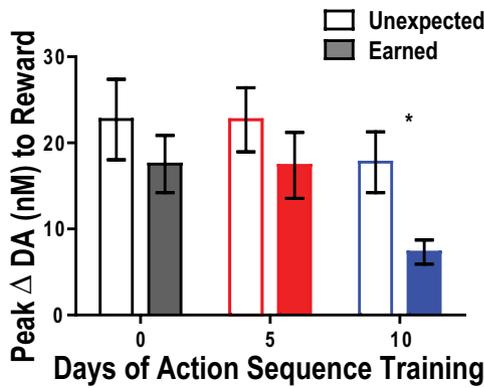
**Figure 5.** Phasic mesolimbic dopamine release backpropagates from reward delivery to preceding the distal action across training groups. **(A)** Change in dopamine concentration in the 5 seconds before and after the distal lever press, averaged across all 30 trials in the test for each rat and then averaged across rats, for rats receiving either 0 (top, black), 5 (middle, red), or 10 (bottom, blue) days of action sequence training before test. Tick marks on each graph represent the average time at which the dopamine peak occurred for each rat. The shaded bar after the distal press line indicates the time of the proximal lever insertion cue. **(B)** Change in dopamine concentration in the 5 seconds before and after the proximal lever press for rats receiving either 0, 5, or 10 days of action sequence training before test. The shaded bar after the distal press line indicates the time from lever press to sucrose pellet delivery. Scale bar marks a 5 nmol/L dopamine concentration change. **(C)** The maximal dopamine peak concentration change preceding the distal action, to the lever insertion cue, preceding the proximal action, and to the reward delivery averaged across all trials for each rat and then averaged across rats in each training group receiving either 0 (gray), 5 (red), or 10 (blue) days of action sequence training before test. Dashed lines and error bars indicate  $\pm 1$  standard error of the mean. 0-day group,  $n = 11$ ; 5-day group,  $n = 11$ ; 10-day group,  $n = 8$ . \* $p < .05$ . DA, dopamine.

for the first, fifth, or tenth day of action sequence training. As Figure 8A demonstrates, we found that flupenthixol administered on test altered the time it took rats to complete the action sequence; there was a main effect of both drug [ $F(1,28) = 8.58, p = .01$ ] and training [ $F(2,28) = 11.24, p = .0003$ ], as well as an interaction between these factors [ $F(2,28) = 7.83, p = .002$ ]. Post hoc analysis found that rats treated with flupenthixol took significantly longer to complete each action sequence, on average, relative to vehicle-treated control rats when administered on the initial day of action sequence learning ( $p < .001$ ) but not when administered after 5 or 10 days of training ( $p > .05$ ). Thus, this aspect of action sequence performance appears to be dopamine-dependent only during initial acquisition of the task, presumably because the underlying learning process requires dopamine signaling. However, flupenthixol also had a more persistent task performance effect, reducing the rate of responding on both levers irrespective of training (Figure 8B and 8C). For both lever-pressing actions,

there was a significant main effect of training [distal:  $F(2,28) = 13.52, p < .0001$ ; proximal:  $F(2,28) = 22.38, p < .0001$ ], as well as a main effect of drug [distal:  $F(1,28) = 30.66, p < .0001$ ; proximal:  $F(1,28) = 18.46, p = .0007$ ], with no significant interaction [distal:  $F(2,28) = 1.50, p = .24$ ; proximal:  $F(2,28) = 1.43, p = .26$ ], indicating that, with training, rats pressed faster on both levers and that flupenthixol attenuated press rate in a manner that did not depend on training history. Taken together, these results confirm that dopamine signaling plays a critical role in the acquisition and performance of self-initiated sequential actions.

## Discussion

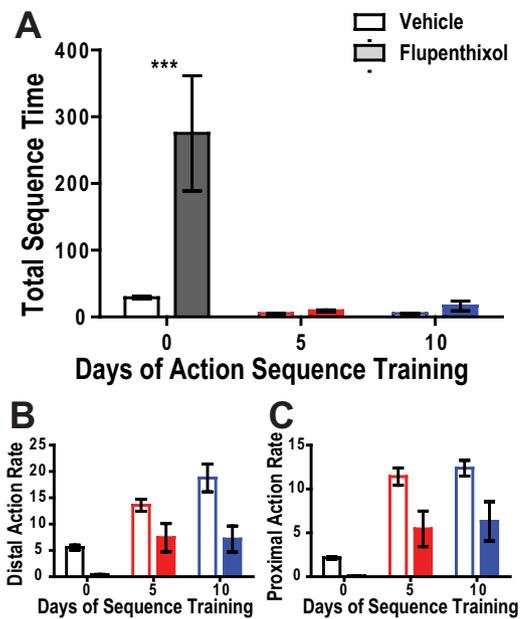
This study characterized the pattern of phasic mesolimbic dopamine release during the acquisition and performance of a self-paced two-action sequencing task in rats. We found that dopamine



**Figure 6.** Dopamine peak amplitude to unexpected and earned reward delivery. Before the lever-pressing session at test, rats were given three free reward deliveries. The dopamine release induced by these unexpected reward deliveries is compared with that induced by delivery of a reward earned by completing the action sequence. Gray bars represent rats with no training on the full action sequence before test, red bars show rats with 5 days of sequence training before test, and blue bars present rats with 10 days of sequence training before test. Error bars indicate  $\pm 1$  standard error of the mean. \* $p < .05$ . DA, dopamine.

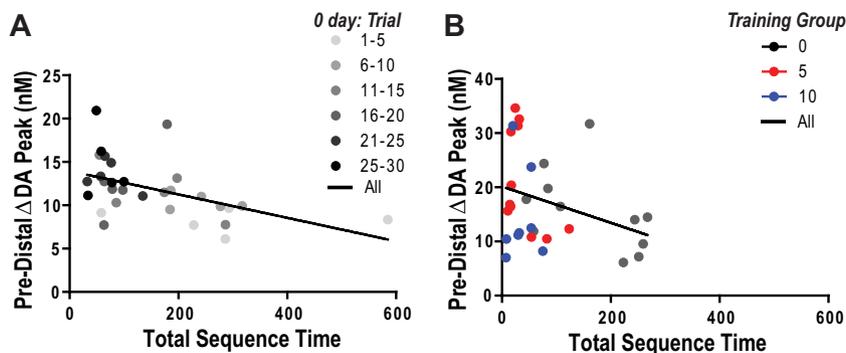
release shifted from the reward to more distal elements of the sequence, a pattern detected both within subjects, in rats acquiring the action sequence for the first time, and across groups of rats given varying amounts of pretraining on the task. Moreover, we found that the concentration of the dopamine transient preceding the initiation of the action sequence predicted the speed with which rats completed the sequence.

These results are generally consistent with the findings of previous studies showing that phasic mesolimbic dopamine signaling transitions from reward delivery to reward-predictive cues during passive Pavlovian learning (7,13,14,41) and discrete-trial, discriminative stimulus-controlled instrumental tasks (10,11,15,42). Our current findings significantly extend this work by demonstrating that phasic dopamine release shifts from reward delivery to precede a self-initiated free-operant instrumental action. In this respect, our results are consistent with the few studies that have examined phasic mesolimbic dopamine release during free-operant behavior for drug reward (31,32,43) and importantly show that phasic mesolimbic dopamine signaling comes to precede a self-



**Figure 8.** Dopamine receptor blockade attenuates action sequence performance. A separate group of rats ( $n = 16$ ) were trained on the action sequence and tested after 0 (gray), 5 (red), or 10 (blue) days of action sequence training. At test, rats were either given saline vehicle (open bars) or flupenthixol (.5 mg/kg intraperitoneal, shaded bars). **(A)** Effects of flupenthixol on action sequence performance, measured by the total action sequence time (averaged across trials then averaged across rats). **(B)** The average rate of lever pressing on the distal lever at test. **(C)** The average rate of lever pressing on the proximal lever at test. Error bars represent  $\pm 1$  standard error of the mean. \*\*\* $p < .0001$ .

initiated action under naturalistic conditions in which the dopamine system is not pharmacologically altered (44,45). Importantly, unlike these studies with drug reward, we show that the phasic dopamine signal to earned food reward diminishes with training. Moreover, our results add to previous studies to suggest that, with learning, the initiation of a self-paced action sequence can be preceded by phasic mesolimbic dopamine release and that performance of an action that has never directly earned reward can be accompanied by phasic dopamine signaling. Interestingly, we show that dopamine release does not transition immediately in our



**Figure 7.** Phasic mesolimbic dopamine release is associated with action sequence performance. **(A)** For each of the 30 trials, the maximum dopamine concentration peak before the distal lever press was averaged across rats (y axis) and correlated with the average time to complete each action sequence performance (totaling the time from initial distal press to the next initiating distal press) and also averaged across rats for each trial (x axis). Shading represents the trial in five-trial bins (light to dark = beginning to end of test). The black line represents the linear regression for all the trials combined. **(B)** For each rat, the dopamine concentration change peak amplitude before the distal lever press was averaged across trials (y axis) and correlated with the average sequence time and also averaged across the 30 trials for each rat (x axis). Gray circles represent rats receiving no action sequence training before test; red and blue circles represent rats receiving 5 and 10 days of action sequence training, respectively. The black bar represents the linear regression for all rats. DA, dopamine.

self-paced action sequence task from the reward to more distal sequence elements but rather that there is an interim learning phase in which phasic dopamine is elevated to each event in the sequence, including the reward delivery. This result is consistent with electrophysiological recordings of midbrain dopamine neurons during cue-reward pairings (41).

Temporal difference models of reinforcement learning assume that learning is regulated by a reward prediction error signal (21,22,46–48), and there is now considerable evidence that this reward prediction error is mediated by phasic dopamine (7,19). However, using a temporal difference algorithm to model free operant conditioning is challenging due to the lack of unambiguous rules for defining model states. This clearly applies to the current task, since our rats were allowed to decide when to initiate the sequence. The main features of our data are nevertheless in line with the general themes of such models. We show that rewards earned by the performance of well-established self-initiated actions elicit significantly smaller phasic mesolimbic dopamine responses than unexpected reward deliveries. Whereas training attenuated the dopamine response to response-contingent rewards, it led to an increase in phasic dopamine release during the period before the distal lever press. This increase was apparent within a single learning session. While not a response to any overt cue, this dopamine response may have been elicited by environmental or internal cues unappreciated by the experimenter.

In addition to the view that dopamine mediates reinforcement learning, a second popular hypothesis assumes that dopamine is responsible for mediating the incentive motivation that allows reward-paired cues to invigorate reward-seeking actions (25,49–51). There have been several recent attempts to integrate the concept that dopamine mediates incentive motivation into the reinforcement learning framework (52–54). For example, McClure *et al.* (52) posit that phasic dopamine's role in mediating the direct incentive motivation effects of reward-predictive cues on action selection is dissociable from its role in reporting the reward prediction errors that support reinforcement learning (22,46,47) or action chunking (55,56). While not providing a critical test of such theories, our finding that the transient dopamine release amplitude preceding the initiation of the action sequence predicts the speed with which that sequence will be completed is generally consistent with a role for phasic mesolimbic dopamine signaling in incentive motivation. Our pharmacological data support this correlational finding by showing that dopamine receptor blockade attenuates both action sequence learning and performance. These effects of dopamine receptor antagonism on action sequence performance are consistent with findings from nonhuman primate studies showing that dopamine transmission is preferentially involved during the early stages of action sequence learning (55,56), when discrete actions are being integrated into sequence-level action chunks (57,58), and with literature implicating dopamine in incentive motivation (25,36,37).

Taken together, these data suggest that phasic mesolimbic dopamine release reflects the properties of a prediction error signal during the acquisition and performance of a self-paced sequence of actions and that such release is also associated with the incentive motivational properties of rewards and reward-paired cues, potentially providing a mechanism by which rewards come to exert influence over temporally distal actions.

*This research was supported by Grants DA09359 and DA05010 from the National Institute on Drug Abuse to NTM, Grant T32 DA024635 from the National Institute on Drug Abuse and Hatos Scholarship to KMW, and Grant DA029035 to SBO.*

*We thank Katie McNutt for research assistance.*

*All authors report no biomedical financial interests or potential conflicts of interest.*

*Supplementary material cited in this article is available online.*

- Hull C (1943): *Principles of Behavior*. New York: Appleton.
- Minsky M (1961): Steps toward artificial intelligence. *Proceedings of the IRE* 49:8–30.
- Jankovic J (2008): Parkinson's disease: Clinical features and diagnosis. *J Neurol Neurosurg Psychiatry* 79:368–376.
- Hyman SE (2007): The neurobiology of addiction: Implications for voluntary control of behavior. *Am J Bioeth* 7:8–11.
- Volkow ND, Fowler JS, Wang GJ, Baler R, Telang F (2009): Imaging dopamine's role in drug abuse and addiction. *Neuropharmacology* 56(suppl 1):3–8.
- Koob GF, Volkow ND (2010): Neurocircuitry of addiction. *Neuropsychopharmacology* 35:217–238.
- Schultz W, Dayan P, Montague PR (1997): A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Schultz W (1999): The reward signal of midbrain dopamine neurons. *News Physiol Sci* 14:249–255.
- Schultz W (2002): Getting formal with dopamine and reward. *Neuron* 36:241–263.
- Ljungberg T, Apicella P, Schultz W (1992): Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol* 67:145–163.
- Schultz W, Apicella P, Ljungberg T, Romo R, Scarnati E (1993): Reward-related activity in the monkey striatum and substantia nigra. *Prog Brain Res* 99:227–235.
- Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010): Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron* 68:815–834.
- Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuh A, Willuhn I, *et al.* (2011): A selective role for dopamine in stimulus-reward learning. *Nature* 469:53–57.
- Day JJ, Roitman MF, Wightman RM, Carelli RM (2007): Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci* 10:1020–1028.
- Roitman MF, Stuber GD, Phillips PE, Wightman RM, Carelli RM (2004): Dopamine operates as a subsecond modulator of food seeking. *J Neurosci* 24:1265–1271.
- Jones JL, Day JJ, Aragona BJ, Wheeler RA, Wightman RM, Carelli RM (2010): Basolateral amygdala modulates terminal dopamine release in the nucleus accumbens and conditioned responding. *Biol Psychiatry* 67:737–744.
- Takikawa Y, Kawagoe R, Hikosaka O (2004): A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward mapping. *J Neurophysiol* 92:2520–2529.
- Schultz W, Apicella P, Ljungberg T (1993): Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J Neurosci* 13:900–913.
- Waelti P, Dickinson A, Schultz W (2001): Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412:43–48.
- Montague PR, Dayan P, Sejnowski TJ (1996): A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947.
- Suri RE, Schultz W (1998): Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Exp Brain Res* 121:350–354.
- Joel D, Niv Y, Ruppel E (2002): Actor-critic models of the basal ganglia: New anatomical and computational perspectives. *Neural Netw* 15:535–547.
- Veeneman MM, van Ast M, Broekhoven MH, Limpens JH, Vanderschuren LJ (2011): Seeking-taking chain schedules of cocaine and sucrose self-administration: Effects of reward size, reward omission, and  $\alpha$ -flupenthixol [published online ahead of print October 12]. *Psychopharmacology (Berl)*.
- Salamone JD, Wisniecki A, Carlson BB, Correa M (2001): Nucleus accumbens dopamine depletions make animals highly sensitive to high fixed ratio requirements but do not impair primary food reinforcement. *Neuroscience* 105:863–870.

25. Berridge KC, Robinson TE (1998): What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev* 28:309–369.
26. Dickinson A, Smith J, Mirenowicz J (2000): Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. *Behav Neurosci* 114:468–483.
27. Wassum KM, Ostlund SB, Balleine BW, Maidment NT (2011): Differential dependence of Pavlovian incentive motivation and instrumental incentive learning processes on dopamine signaling. *Learn Mem* 18:475–483.
28. Ostlund SB, Maidment NT (2012): Dopamine receptor blockade attenuates the general incentive motivational effects of noncontingently delivered rewards and reward-paired cues without affecting their ability to bias action selection. *Neuropsychopharmacology* 37:408–519.
29. Romo R, Schultz W (1990): Dopamine neurons of the monkey midbrain: Contingencies of responses to active touch during self-initiated arm movements. *J Neurophysiol* 63:592–606.
30. Jin X, Costa RM (2010): Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* 466:457–462.
31. Stuber GD, Roitman MF, Phillips PE, Carelli RM, Wightman RM (2005): Rapid dopamine signaling in the nucleus accumbens during contingent and noncontingent cocaine administration. *Neuropsychopharmacology* 30:853–863.
32. Phillips PE, Stuber GD, Heien ML, Wightman RM, Carelli RM (2003): Subsecond dopamine release promotes cocaine seeking. *Nature* 422:614–618.
33. Redish AD (2004): Addiction as a computational process gone awry. *Science* 306:1944–1947.
34. Anderson SM, Pierce RC (2005): Cocaine-induced alterations in dopamine receptor signaling: Implications for reinforcement and reinstatement. *Pharmacol Ther* 106:389–403.
35. Romo R, Schultz W (1989): Somatosensory input to dopamine neurones of the monkey midbrain: Responses to pain pinch under anaesthesia and to active touch in behavioural context. *Prog Brain Res* 80:473–478; discussion 465–466.
36. Robbins TW, Everitt BJ (2007): A role for mesencephalic dopamine in activation: Commentary on Berridge (2006). *Psychopharmacology (Berl)* 191:433–437.
37. Salamone JD, Correa M, Farrar A, Mingote SM (2007): Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl)* 191:461–482.
38. Ostlund SB, Wassum KM, Murphy NP, Balleine BW, Maidment NT (2011): Extracellular dopamine levels in striatal subregions track shifts in motivation and response cost during instrumental conditioning. *J Neurosci* 31:200–207.
39. Lynd-Balta E, Haber SN (1994): The organization of midbrain projections to the striatum in the primate: Sensorimotor-related striatum versus ventral striatum. *Neuroscience* 59:625–640.
40. Lynd-Balta E, Haber SN (1994): The organization of midbrain projections to the ventral striatum in the primate. *Neuroscience* 59:609–623.
41. Pan WX, Schmidt R, Wickens JR, Hyland BI (2005): Dopamine cells respond to predicted events during classical conditioning: Evidence for eligibility traces in the reward-learning network. *J Neurosci* 25:6235–6242.
42. Nishino H, Ono T, Muramoto K, Fukuda M, Sasaki K (1987): Neuronal activity in the ventral tegmental area (VTA) during motivated bar press feeding in the monkey. *Brain Res* 413:302–313.
43. Owesson-White CA, Ariansen J, Stuber GD, Cleaveland NA, Cheer JF, Wightman RM, Carelli RM (2009): Neural encoding of cocaine-seeking behavior is coincident with phasic dopamine release in the accumbens core and shell. *Eur J Neurosci* 30:1117–1127.
44. Addy NA, Daberkow DP, Ford JN, Garris PA, Wightman RM (2010): Sensitization of rapid dopamine signaling in the nucleus accumbens core and shell after repeated cocaine in rats. *J Neurophysiol* 104:922–931.
45. Whitby LG, Hertting G, Axelrod J (1960): Effect of cocaine on the disposition of noradrenaline labelled with tritium. *Nature* 187:604–605.
46. Sutton RS (1988): Learning to predict by methods of temporal differences. *Machine Learning* 3:9–44.
47. Barto AG, Sutton RS, Watkins CJCH (1989): Sequential decision problems and neural networks. In: Touretzky DS, editor. *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 686–693.
48. Houk JC, Adams JL, Barto AG (1995): A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: Houk JC, Davis JL, Beiser DG, editors. *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press, 249–270.
49. Berridge KC (2007): The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology (Berl)* 191:391–431.
50. Berridge KC, Robinson TE, Aldridge JW (2009): Dissecting components of reward: 'Liking', 'wanting', and learning. *Curr Opin Pharmacol* 9:65–73.
51. Ikemoto S, Panksepp J (1999): The role of nucleus accumbens dopamine in motivated behavior: A unifying interpretation with special reference to reward-seeking. *Brain Res Brain Res Rev* 31:6–41.
52. McClure SM, Daw ND, Montague PR (2003): A computational substrate for incentive salience. *Trends Neurosci* 26:423–428.
53. Zhang J, Berridge KC, Tindell AJ, Smith KS, Aldridge JW (2009): A neural computational model of incentive salience. *PLoS Comput Biol* 5:e1000437.
54. Niv Y, Daw ND, Joel D, Dayan P (2007): Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191:507–520.
55. Tremblay PL, Bedard MA, Levesque M, Chebli M, Parent M, Courtemanche R, Blanchet PJ (2009): Motor sequence learning in primate: Role of the D2 receptor in movement chunking during consolidation. *Behav Brain Res* 198:231–239.
56. Levesque M, Bedard MA, Courtemanche R, Tremblay PL, Scherzer P, Blanchet PJ (2007): Raclopride-induced motor consolidation impairment in primates: Role of the dopamine type-2 receptor in movement chunking into integrated sequences. *Exp Brain Res* 182:499–508.
57. Ostlund SB, Winterbauer NE, Balleine BW (2009): Evidence of action sequence chunking in goal-directed instrumental conditioning and its dependence on the dorsomedial prefrontal cortex. *J Neurosci* 29:8280–8287.
58. Graybiel AM (1998): The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem* 70:119–136.
59. Paxinos G, Watson C (1998): *The Rat Brain in Stereotaxic Coordinates, 4th ed.* San Diego: Academic Press.