Brief Communications

Dopamine Encoding of Pavlovian Incentive Stimuli Diminishes with Extended Training

Jeremy J. Clark,[1] **Anne L. Collins,**[1] **Christina Akers Sanford,**[1,2] **and Paul E. M. Phillips**[1,2]

Departments of [1]Psychiatry and Behavioral Sciences, and [2]Pharmacology, University of Washington, Seattle, Washington 98195

Dopamine is highly implicated both as a teaching signal in reinforcement learning and in motivating actions to obtain rewards. However, theoretical disconnects remain between the temporal encoding properties of dopamine neurons and the behavioral consequences of its release. Here, we demonstrate in rats that dopamine evoked by pavlovian cues increases during acquisition, but dissociates from stable conditioned appetitive behavior as this signal returns to preconditioning levels with extended training. Experimental manipulation of the statistical parameters of the behavioral paradigm revealed that this attenuation of cue-evoked dopamine release during the postasymptotic period was attributable to acquired knowledge of the temporal structure of the task. In parallel, conditioned behavior became less dopamine dependent after extended training. Thus, the current work demonstrates that as the presentation of reward-predictive stimuli becomes anticipated through the acquisition of task information, there is a shift in the neurobiological substrates that mediate the motivational properties of these incentive stimuli.

Introduction

Reward-related dopamine transmission within the mesolimbic system is hypothesized to function as a reinforcement signal that promotes future behavioral responses to predictive cues (Wise, 2004) as well as a motivational signal that immediately mobilizes behavior through the assignment of incentive value (Berridge, 2007). Phasic dopamine neurotransmission during the contingent pairing of conditioned stimuli (CS) and rewards [unconditioned stimuli (US)] shows a dynamic pattern of signaling where US-evoked phasic responses gradually decrease in parallel with a gradual increase in CS-evoked responses (Ljungberg et al., 1992). This pattern is highly relevant to the motivational properties of stimuli as it is differentially regulated dependent upon the degree to which individuals assign incentive value to reward-predictive cues (Flagel et al., 2011). Indeed, acquired phasic dopamine release at the time of a CS may function similarly to that of primary rewards to provide conditional reinforcement supporting secondary conditioning through the assignment of incentive value (McClure et al., 2003).

In the context of reinforcement learning, decreased US-evoked dopamine during learning is attributed to its developing predictability by the presentation of the CS, and increased CS-evoked dopamine is attributed to the establishment of this stimulus as the earliest predictor of reward. Thus, the presence of a dopamine signal only when rewards are not fully predicted is interpreted as evidence for dopamine acting as a teaching signal to update predictions when they are not accurate with regard to the precise timing and value of impending reward. Consistent with a significant role for predictability, CS-evoked responses also diminish when preceded by cues that occur at regular time intervals (Schultz, 1998), confirming that timing of reward-related events is central to the generation of these signals (Fiorillo et al., 2008) as well as a critical component to learning (Gallistel and Gibbon, 2000).

However, many real-world situations involve uncertainty in the probability and/or timing of rewards and reward-predictive cues. In experimental paradigms involving probabilistic rewards, there is evidence that cue-evoked dopamine signaling scales with the probability of reward delivery (Fiorillo et al., 2003). The generation of anticipatory behavior such as approach requires not only knowledge concerning the variability in rewarding outcomes but also the ability to track the temporal pattern of cues for estimating the likelihood of an event occurring at a given time (i.e., hazard rate; Janssen and Shadlen, 2005). Thus, knowledge of task statistics, perhaps acquired through extended experience, may modulate dopamine-encoded prediction errors. However, the evolution of such responses during learning and the environmental conditions contributing to their development remain unclear. In addition to questions regarding the temporal encoding properties of dopamine neurons, these concepts also highlight a theoretical disconnect between the environmental events encoded by dopamine neurotransmission and the behavioral consequences of dopamine release. Indeed, if stimulus-evoked (CS or US) phasic dopamine transmission is attenuated as stimuli become predicted, it is unclear whether and how the motivational properties of these stimuli are transmitted and maintained.

Materials and Methods

Animals. Male Sprague Dawley rats weighing ∼300–350 g were obtained from Charles River, were housed individually on a 12 h light/dark cycle

Figure 1. Conditioned approach behavior throughout acquisition and postasymptotic training. ***A***, Probability to lever press by training session. ***B–D***, Probability to lever press (***B***), total lever presses (***C***), and latency to lever press (***D***) binned by decades of trials. The green line depicts the best fit from the Weibull function for each behavioral measure. Vertical dotted lines denote the decade in training where asymptotic performance was reached, as determined by the best fit parameter from the Weibull function. Data are mean ± SEM.
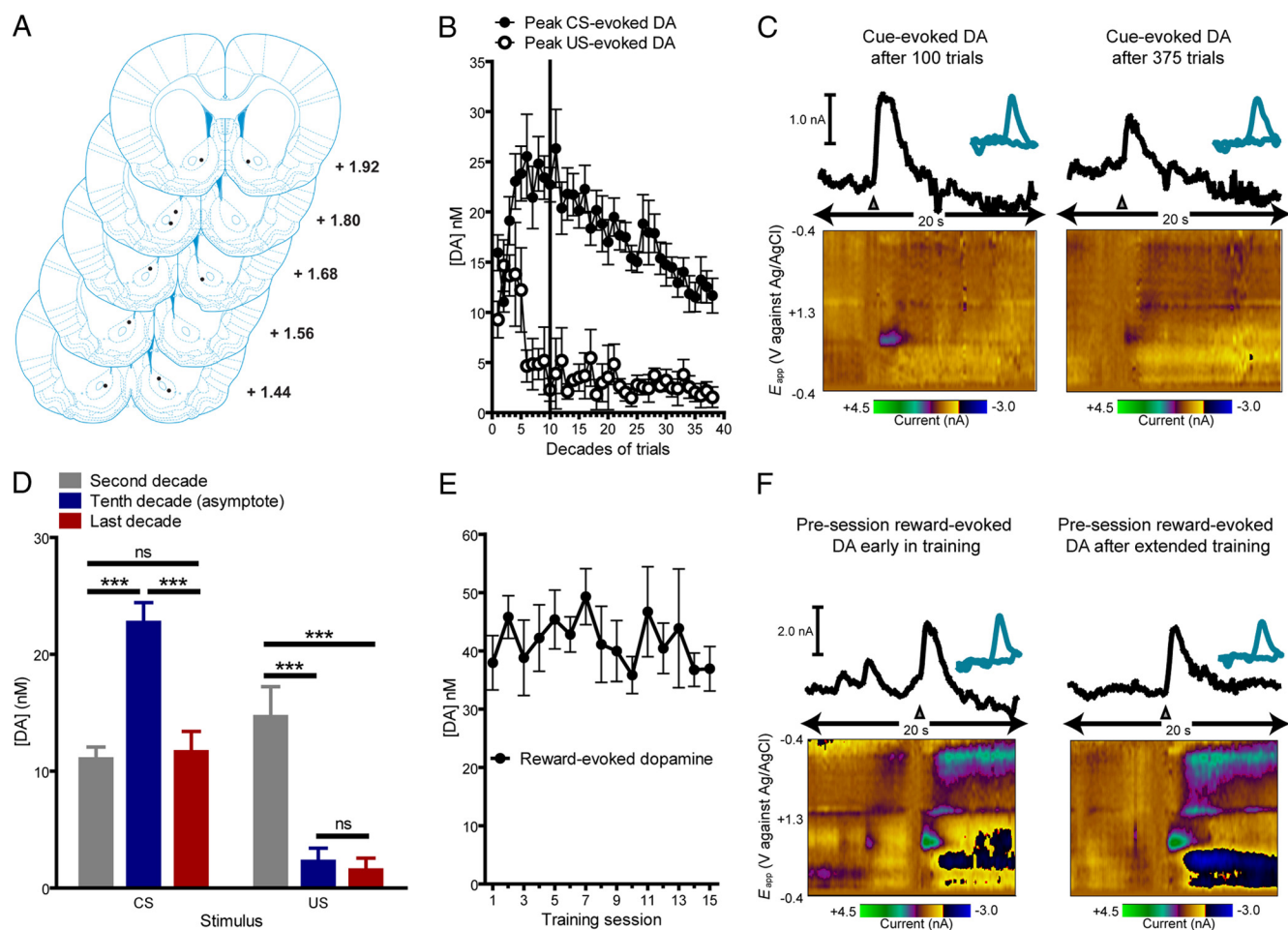
average postimplantation electrode sensitivity (Clark et al., 2010). Peak CS- and US-evoked dopamine release values were obtained by taking the largest value in the 2 s period after stimulus presentation. Mixed-measures ANOVA was used to compare peak stimulus-evoked dopamine release during learning with stimulus as the between-group measure and decades as the within-group measure. Separate repeated-measures ANOVA for CS, US, and presession rewards were used to assess stimulus-evoked dopamine release across both phases of training with *post hoc* tests for linear trends. CS- and US-evoked dopamine release during the first, 10th, and last decade was compared with two-way ANOVA and *post hoc* t tests with the Bonferroni correction for multiple tests.

Behavior. Following a single session of magazine training where 20 food pellets (45 mg; Bio-Serve) were delivered at a 90 s variable interval, rats were trained on a pavlovian conditioned approach task (Flagel et al., 2011). During daily sessions, 25 trials were presented with a variable intertrial interval (ITI) from a range of values consisting of 30, 40, 50, 60, 70, 80, and 90 s (without replacement). A trial consisted of a lever/light cue presented for 8 s followed immediately by delivery of a food pellet and retraction of the lever. Lever presses were recorded but without consequence for reward delivery. Animals failing to approach the predictive cue by the fifth session on at least 75% of trials, as measured by lever pressing, were excluded from subsequent analysis (n = 10). This criterion selects rats that approach the predictive cue (sign tracking) and excludes animals that approach the site of reward delivery during cue presentation (goal tracking) as these behaviors are differentially dependent upon intact dopamine neurotransmission (Flagel et al., 2011) and may reflect different learning mechanisms (Clark et al., 2012). Behavioral data were binned into 10-trial epochs and fit with a standard psychometric function (Weibull function) to obtain the best fit parameter for asymptote. Conditioned approach behavior was compared with stimulus-evoked dopamine release with linear regression separately for the preasymptotic and postasymptotic phases. All statistical analyses were performed using Prism (GraphPad Software).

Probe trials. After 15 sessions (375 trials), all animals were given two sessions that included probe trials, counterbalanced for order of presentation and separated by one normal session of training. In each probe session, 5 probe trials were presented along with 20 standard trials. For CS probe trials, 5 trials were presented with an ITI of 10 s, with the remaining 20 trials occurring within the normal range of ITI values. For US probe trials, all trials were identical to normal training sessions with the exception that an uncued reward was delivered during the ITI after every fifth trial. For CS probes, paired t tests were used to compare cue-evoked dopamine release on probe trials to cue-evoked dopamine release on normal trials within the same session. Independent-sample t tests were used for comparison of cue-evoked dopamine release on short (<60 s) to that of long (>60 s) ITI values. For US probes, paired t tests were used to compare reward-evoked dopamine release on probe trials to reward-evoked dopamine release on normal trials within the same session.

Histological verification of recording sites. Animals were anesthetized with sodium pentobarbital; the recording site was then marked with an electrolytic lesion (300 V) by applying current directly through the re-

with Teklad rodent chow and water available *ad libitum* except as noted, and were weighed and handled daily. Before conditioning tasks, rats were food deprived to ~90% of their free-feeding body weight. All experimental procedures were in accordance with the Institutional Animal Care and Use Committee at the University of Washington.

Surgery and electrochemical detection of dopamine. Rats (n = 30) were implanted with carbon-fiber microelectrodes (1.3 mm lateral, 1.3 mm rostral, and 6.8 mm ventral of bregma) for *in vivo* detection of phasic dopamine using fast-scan cyclic voltammetry (Clark et al., 2010). Thirty minutes before the start of each experimental session, rats were placed in an operant chamber (Med Associates) and chronically implanted microsensors were connected to a head-mounted voltammetric amplifier. Of the 20 animals meeting the behavioral criterion, 3 had electrode placements outside of the nucleus accumbens core and 7 failed for technical reasons (e.g., loss of headcap, saturation of signal). Rats (n = 10) were given a single uncued food pellet, delivered to the food receptacle, before the start of each session to assess reward-evoked dopamine signaling. Voltammetric scans were repeated every 100 ms (−0.4 to +1.3 V at 400 V/s; National Instruments), and dopamine was isolated from the voltammetric signal with chemometric analysis (Heien et al., 2005) using a standard training set based on stimulated dopamine release detected by chronically implanted electrodes. Dopamine concentration was estimated based on the

Figure 2. Dopamine dynamics during preasymptotic and postasymptotic pavlovian conditioning. *A*, All recording sites (●) were within the nucleus accumbens core. The numbers on each plate indicate the distance in millimeters anterior from bregma (Paxinos and Watson, 2005). *B*, Average peak CS- and US-evoked dopamine release across 375 trials of conditioning. *C*, Color-coded observed changes in electrochemical information as a function of applied potential (*y*-axis) plotted over time (*x*-axis). *D*, Comparison of peak US-evoked and CS-evoked phasic dopamine release at different points in training. *E*, Reward-evoked dopamine release before the start of each training session. Reward-evoked dopamine release did not significantly change over sessions as determined by one-way repeated-measures ANOVA ($F_{(14,126)}$ = 0.79, $p > 0.05$; post-test for linear trend, $p > 0.05$). *F*, Color-coded observed changes in electrochemical information as a function of applied potential (*y*-axis) plotted over time (*x*-axis). Gray triangles denote CS onset or reward delivery. Data are mean ± SEM. ***$p < 0.001$. ns, Not a statistically significant difference.

cording electrode for 20 s. Animals were transcardially perfused with PBS followed by 4% paraformaldeyde. Brains were removed, and after were fixed in paraformaldehyde and then rapidly frozen in an isopentane bath (~5 min), sliced on a cryostat (50 μm coronal sections, 20°C), and stained with cresyl violet to aid in visualization of anatomical structures.

Pharmacology. A separate cohort of rats was trained as above on a pavlovian conditioned approach task ($n = 40$) for either 5 sessions (asymptotic training group) or 15 sessions (postasymptotic training group). Fourteen rats failed to reach a criterion of 75% approach by the fifth session and were excluded from analysis. The last session of training, either the fifth or 15th, was followed by a test session where animals received five cue presentations in extinction. Thirty minutes before the test session, animals were injected with either the dopamine D1 receptor antagonist SCH23390 (0.01 mg/kg, i.p.) or saline. Two-way ANOVA with training and drug condition as between-group measures was used to assess conditioned approach behavior on the test day followed by *post hoc* t tests with the Bonferroni correction for multiple tests.

Results

Over 15 sessions (375 trials), we observed conditioned approach behavior that increased over the first 4 sessions and remained stable thereafter ($n = 10$; Fig. 1A). To determine asymptotic performance level, we analyzed three separate behavioral measures in 10-trial epochs (Fig. 1B–D) and fit these data with the Weibull

function, a standard psychometric tool in the analysis of learning curves (Gallistel et al., 2004). The time to reach asymptote was defined as the first decade where mean response level exceeded the 95% confidence interval (CI) of the best fit parameter for asymptote from each behavioral measure (asymptote for probability = 0.87, 95% CI = 0.83−0.90; total lever presses = 28.69, 95% CI = 26.85−30.53; latency = 3.55, 95% CI = 3.36−3.74). We used this statistic, similar across all behavioral metrics (Fig. 1B–D), to divide behavior into preasymptotic (100 trials) and postasymptotic periods for neurochemical analysis (Fig. 2A,B). During the preasymptotic period (first 100 trials), there was a trial-by-trial shift in phasic dopamine activity from the reward to the CS, in agreement with previous reports (Flagel et al., 2011). Consistent with the encoding of a reward prediction error, cue-evoked phasic dopamine increased ($F_{(9,81)} = 6.14$, $p < 0.0001$; post-test for linear trend, $p < 0.0001$) and was positively correlated ($r^2 = 0.46$, $p < 0.05$) with conditioned approach, while reward-evoked dopamine decreased ($F_{(9,81)} = 4.54$, $p < 0.0001$; post-test for linear trend, $p < 0.0001$) and was negatively correlated ($r^2 = 0.77$, $p < 0.0001$) with conditioned approach (Fig. 2B,C). In the postasymptotic period of training (trials 100−375), the dopamine response to the US did not change further and
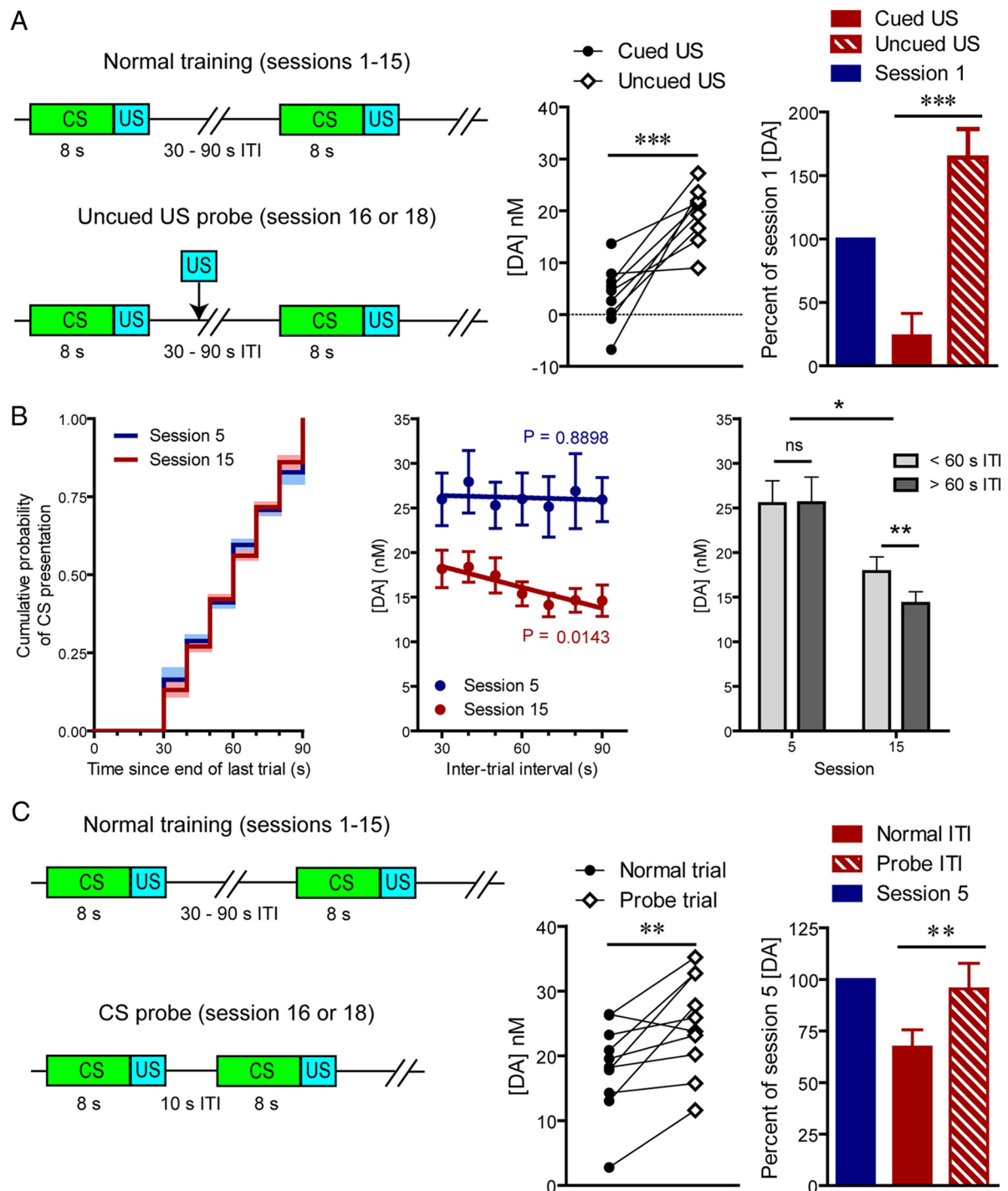
Figure 3. Effect of temporal expectation on stimulus-evoked dopamine release. **A**, Illustration of task design (left) and US-evoked dopamine during test sessions (middle) relative to preconditioning level (right). **B**, Experienced cumulative probability of CS presentation during sessions 5 and 15 (left), CS-evoked dopamine as a function of ITI length (middle), and comparison of short to long ITIs at postacquisition asymptote and after extended training (right). **C**, Illustration of task design (left) and CS-evoked dopamine during the probe session (middle) relative to peak levels at postacquisition asymptote (right). Data are mean ± SEM. $*p < 0.05$, $**p < 0.01$, $***p < 0.005$.

remained minimal throughout this period ($F_{(9,234)} = 0.58$, $p > 0.05$; post-test for linear trend, $p > 0.05$). However, cue-evoked dopamine release declined during the postasymptotic period back to preconditioning levels ($F_{(9,234)} = 4.46$, $p < 0.0001$; post-test for linear trend, $p < 0.0001$). Comparison of peak US-evoked and CS-evoked phasic dopamine release at different points in training (Fig. 2D) using mixed measures ANOVA with stimulus (CS and US) as the between-group measure and decade of train-

ing (second, fifth, and 10th) as the within-group measure revealed a significant main effect of stimulus ($F_{(1,52)} = 44.85$, $p < 0.0001$), a significant main effect of decade ($F_{(2,52)} = 9.13$, $p < 0.005$), and a significant interaction effect between stimulus and decade ($F_{(2,52)} = 26.95$, $p < 0.0001$). *Post hoc* tests showed that US-evoked dopamine release was significantly lower on the 10th ($p < 0.001$) and last decade ($p < 0.0001$) of training compared with the second decade. However, the 10th and last decades were

not significantly different from each other. Conversely, *post hoc* tests revealed that CS-evoked dopamine release increased significantly from the second to the 10th decade ($p < 0.001$) and then significantly decreased back to preconditioning levels from the 10th to the last decade ($p < 0.001$), where it did not differ from the preconditioning level.

Stable reward-evoked dopamine release observed outside the context of the task (Fig. 2 E, F) indicates that attenuation of cue-evoked signaling during the postasymptotic period is not attributable to general degradation of dopamine transmission. Therefore, we tested whether there was a development of task-related contextual suppression of dopamine release over the course of training (Fig. 3A). Not surprisingly, when uncued rewards were delivered during the task (session 16 or 18), they elicited significantly more dopamine release than cued rewards ($t_{(9)} = 5.42$, $p < 0.001$; Fig. 3A). Importantly, the level of dopamine release to uncued rewards during this postasymptotic phase was restored to the preacquisition level (session 1; Fig. 3A), indicating that any contextual suppression did not develop over this period. Having ruled out these possibilities, we hypothesized that attenuation of CS-evoked dopamine release was conferred by the acquisition of a temporal expectation of CS presentation. This notion is somewhat surprising given that CS presentation occurred at variable time intervals with respect to the end of previous trials. Nonetheless, if animals had acquired knowledge about the temporal statistics of the task, we would anticipate that their expectation would correspond to the hazard rate (Fig. 3B) where the shortest time interval would be less predictable than progressively longer ones and, importantly, that this expectation would modulate the magnitude of cue-evoked phasic dopamine. Moreover, this temporal estimation would be expected to develop after the cue becomes established as a full predictor of reward and, as such, should be present after postasymptotic training (session 15) but not immediately after acquisition (session 5). Consistent with our hypothesis, a pattern emerged over the course of extended training where higher cue-evoked dopamine release was observed for shorter ITIs (main effect of ITI: $F_{(1,18)} = 5.14$, $p < 0.05$; main effect of session: $F_{(1,18)} = 9.91$, $p < 0.01$; session × ITI interaction: $F_{(1,18)} = 5.78$, $p < 0.05$), resulting in significant correlation between phasic dopamine signaling and the ITI after 15 sessions ($p < 0.05$) of training but not after 5 sessions ($p > 0.05$; Fig. 3B). Therefore, to further test our hypothesis we conducted probe trials where cues were presented with a shorter ITI than previously experienced by the animals (Fig. 3C). These probe trials elicited significantly higher dopamine release than regular trials ($t_{(9)} = 3.46$, $p < 0.01$; Fig. 3C) and recovered signaling to that of session 5, suggesting that attenuation can be solely attributed to the learning of task statistics.

Stable conditioned approach behavior accompanied by diminishing cue-evoked dopamine release introduces a notable dissociation between a behavioral hallmark of acquired incentive value and dopamine encoding of pavlovian cues (Fig. 4A). This separation suggests that the involvement of dopamine in conditioned behavior may change during postasymptotic learning. Indeed, a diminishing role of dopamine over training has been shown for other reward-related behaviors (Choi et al., 2005). Thus, to determine the dependence of conditioned approach on dopamine D1 receptor activation at postacquisition asymptote and after extended postasymptotic training, animals were trained on the pavlovian conditioned approach task for 125 or 375 trials and then received either the dopamine D1 receptor antagonist SCH23390 or saline during a test session ($n = 26$). Conditioned approach on the last day of training before the test session did not
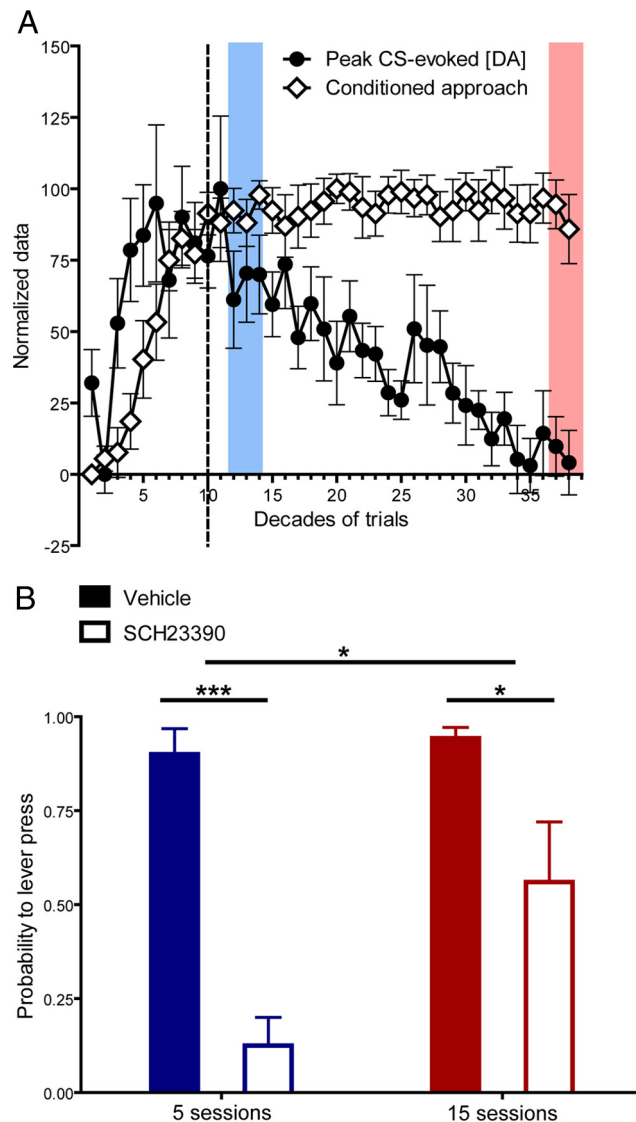
Figure 4. Temporal dynamics of CS-evoked dopamine and dopamine dependence of conditioned behavior. **A**, Normalization of conditioned approach behavior and CS-evoked dopamine across all phases of training. **B**, Conditioned approach behavior on a test session after injection of SCH23390 or saline in animals that received either 5 sessions (saline: $n = 6$; SCH23390: $n = 8$; blue) or 15 sessions (saline: $n = 7$; SCH23390: $n = 5$; red) of training. Shaded blocks in **A** correspond to the point in training where pharmacological experiments were conducted in **B**. The vertical dotted line denotes the decade in training where asymptotic performance was reached. Data are mean ± SEM. *$p < 0.05$, ***$p < 0.005$.

significantly differ between groups. D1 receptor antagonism significantly reduced the conditioned approach for both periods of training but was less effective following extended postasymptotic training (main effect of drug: $F_{(1,22)} = 45.39$, $p < 0.0001$; main effect of training: $F_{(1,22)} = 7.73$, $p < 0.05$; drug × training interaction effect: $F_{(1,22)} = 5.21$, $p < 0.05$; Fig. 4B), demonstrating that the dopamine dependence of conditioned behavior changes during postasymptotic training.

Discussion

A role for dopamine in reinforcement learning is suggested by the correlation between phasic patterns of neurotransmission during the contingent pairing of rewards and predictive stimuli and the encoding of a reward prediction error used as a teaching signal in formal models of learning (Montague et al., 1996). How-

ever, the contribution of dopamine neurotransmission to processes necessary for the acquisition of conditioned responses during learning and those necessary for maintaining the motivational value that drives performance remains unclear. It has been previously demonstrated that dopamine signaling is required for the acquisition and performance of conditioned approach behavior (Di Ciano et al., 2001) generated by the acquired incentive properties of conditioned stimuli. Specifically, signaling at the dopamine D1 receptor has been associated with phasic dopamine release (Dreyer et al., 2010). Therefore, we compared the effects of a selective dopamine D1 receptor antagonist on CS-elicited conditioned behavior early and late in postasymptotic training, when CS-evoked phasic dopamine was at its peak or after attenuation, respectively. We found that performance of conditioned approach behavior was completely abolished by dopamine D1 receptor antagonism administered at behavioral asymptote but became significantly less dependent on intact D1 signaling after extended postasymptotic training. These findings demonstrate that the incentive properties of conditioned stimuli become less dependent upon dopamine following extended training.

One of the defining features of acquired incentive value by a pavlovian cue is the ability to elicit approach behavior despite the fact that engaging the cue has no instrumental consequence to obtaining reward (Berridge, 2007). Here, we examined pavlovian incentive value, which has been theoretically and experimentally distinguished from instrumental incentive value (Dickinson et al., 2000). Previous work with instrumental learning has demonstrated a transition in the underlying associative structure of conditioned behavior across training where early in training responding is sensitive to manipulations of reward outcome but becomes increasingly insensitive as training progresses (i.e., behavior becomes habitual; Dickinson, 1985). This behavioral change is accompanied by a switch in the dopamine dependence of performance from intact dopamine neurotransmission in the ventral striatum to intact dopamine neurotransmission in the dorsal striatum (Vanderschuren et al., 2005). Thus, the current findings demonstrate an important contrast between instrumental and pavlovian conditioning where the switch in the underlying mechanism for conditioned responding is based upon dopamine dependence in different structures for the former and a less dopamine-dependent state in general for the latter.

The observed attenuation of cue-evoked dopamine release after extended pavlovian training mirrors findings of a previous report where the phasic activation of midbrain dopamine neurons in response to cues signaling reward availability was shown to attenuate after extensive overtraining (Ljungberg et al., 1992). Here we show that this attenuation is attributable to the developing predictability of trial onset, comparable to that described for manipulations of CS duration (Fiorillo et al., 2008), as animals learn a hazard rate conferred by the statistical parameters of the task. Indeed, the timing of rewards and their predictors is an integral feature to many theoretical accounts of learning (Savastano and Miller, 1998) and an important contribution of the computational reinforcement learning framework (Sutton and Barto, 1998) to traditional associative models (Rescorla and Wagner, 1972).

An alternative interpretation of attenuated cue-evoked dopamine release is that event predictability can become established through occasion setting where predictive information about stimulus delivery is provided by the context. Occasion setters offer configural information on expected contingencies

between discrete stimuli (Myers and Gluck, 1994). Accordingly, this account would predict that, following sufficient training, CS-US presentation within the context of the session would elicit decreasing phasic dopamine release as the context comes to predict it. However, if the context were suppressing cue-evoked dopamine release after extended training, we would anticipate that this suppression would be present regardless of the temporal relationship between cues (the intertrial interval). Contrary to this prediction, probe trials after extended training presented at shortened time intervals returned cue-evoked dopamine signaling to the preattenuation levels obtained during session 5, supporting the conclusion that attenuation can be attributed to estimates of temporal task statistics and not contextual learning.

These findings provide neurobiological evidence for the encoding of temporal information that could be used to shape and guide adaptive preparatory behavior through the generation of estimates of upcoming events, even if they occur at irregular intervals. Collectively, they demonstrate that dopamine-encoded prediction errors are modulated by ongoing estimates in the timing of reward-predictive events, dissociating them from the motivational significance of these events as they become anticipated.

References

Berridge KC (2007) The debate over dopamine's role in reward: the case for incentive salience. Psychopharmacology 191:391–431. CrossRef Medline

Choi WY, Balsam PD, Horvitz JC (2005) Extended habit training reduces dopamine mediation of appetitive response expression. J Neurosci 25:6729–6733. CrossRef Medline

Clark JJ, Sandberg SG, Wanat MJ, Gan JO, Horne EA, Hart AS, Akers CA, Parker JG, Willuhn I, Martinez V, Evans SB, Stella N, Phillips PE (2010) Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. Nat Methods 7:126–129. CrossRef Medline

Clark JJ, Hollon NG, Phillips PE (2012) Pavlovian valuation systems in learning and decision making. Curr Opin Neurobiol 22:1054–1061. CrossRef Medline

Di Ciano P, Cardinal RN, Cowell RA, Little SJ, Everitt BJ (2001) Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens core in the acquisition and performance of pavlovian approach behavior. J Neurosci 21:9471–9477. Medline

Dickinson A (1985) Actions and habits: the development of behavioural autonomy. Philos Trans R Soc Lond B 308:67–78. CrossRef

Dickinson A, Smith J, Mirenowicz J (2000) Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. Behav Neurosci 114:468–483. CrossRef Medline

Dreyer JK, Herrik KF, Berg RW, Hounsgaard JD (2010) Influence of phasic and tonic dopamine release on receptor activation. J Neurosci 30:14273–14283. CrossRef Medline

Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. Science 299:1898–1902. CrossRef Medline

Fiorillo CD, Newsome WT, Schultz W (2008) The temporal precision of reward prediction in dopamine neurons. Nat Neurosci 11:966–973. CrossRef Medline

Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PE, Akil H (2011) A selective role for dopamine in stimulus-reward learning. Nature 469:53–57. CrossRef Medline

Gallistel CR, Gibbon J (2000) Time, rate, and conditioning. Psychol Rev 107:289–344. CrossRef Medline

Gallistel CR, Fairhurst S, Balsam P (2004) The learning curve: implications of a quantitative analysis. Proc Natl Acad Sci U S A 101:13124–13131. CrossRef Medline

Heien ML, Khan AS, Ariansen JL, Cheer JF, Phillips PE, Wassum KM, Wightman RM (2005) Real-time measurement of dopamine fluctuations after cocaine in the brain of behaving rats. Proc Natl Acad Sci U S A 102:10023–10028. CrossRef Medline

Janssen P, Shadlen MN (2005) A representation of the hazard rate of elapsed time in macaque area LIP. Nat Neurosci 8:234–241. CrossRef Medline

Ljungberg T, Apicella P, Schultz W (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. J Neurophysiol 67:145–163. Medline

McClure SM, Daw ND, Montague PR (2003) A computational substrate for incentive salience. Trends Neurosci 26:423–428. CrossRef Medline

Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16:1936–1947. Medline

Myers CE, Gluck MA (1994) Context, conditioning, and hippocampal re-representation in animal learning. Behav Neurosci 108:835–847. CrossRef Medline

Paxinos G, Watson C (2005) The rat brain in stereotaxic coordinates. Amsterdam: Elsevier Academic.

Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In: Classical conditioning II: current research and theory (Black AH, Prokasy WF, eds), pp 64–99. New York: Appleton-Century-Crofts.

Savastano HI, Miller RR (1998) Time as content in Pavlovian conditioning. Behav Processes 44:147–162. CrossRef

Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27. Medline

Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT Press.

Vanderschuren LJ, Di Ciano P, Everitt BJ (2005) Involvement of the dorsal striatum in cue-controlled cocaine seeking. J Neurosci 25:8665–8670. CrossRef Medline

Wise RA (2004) Dopamine, learning and motivation. Nat Rev Neurosci 5:483–494. CrossRef Medline