Behavioral/Systems/Cognitive

Coding of Predicted Reward Omission by Dopamine Neurons in a Conditioned Inhibition Paradigm

Philippe N. Tobler,[1] **Anthony Dickinson,**[2] **and Wolfram Schultz**[1]

[1]Department of Anatomy, University of Cambridge, Cambridge CB2 3DY, United Kingdom, and Institute of Physiology, University of Fribourg, 1700 Fribourg, Switzerland, and [2]Department of Experimental Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom

Animals learn not only about stimuli that predict reward but also about those that signal the omission of an expected reward. We used a conditioned inhibition paradigm derived from animal learning theory to train a discrimination between a visual stimulus that predicted reward (conditioned excitor) and a second stimulus that predicted the omission of reward (conditioned inhibitor). Performing the discrimination required attention to both the conditioned excitor and the inhibitor; however, dopamine neurons showed very different responses to the two classes of stimuli. Conditioned inhibitors elicited considerable depressions in 48 of 69 neurons (median of 35% below baseline) and minor activations in 29 of 69 neurons (69% above baseline), whereas reward-predicting excitors induced pure activations in all 69 neurons tested (242% above baseline), thereby demonstrating that the neurons discriminated between conditioned stimuli predicting reward versus nonreward. The discriminative responses to stimuli with differential reward-predicting but common attentional functions indicate differential neural coding of reward prediction and attention. The neuronal responses appear to reflect reward prediction errors, thus suggesting an extension of the correspondence between learning theory and activity of single dopamine neurons to the prediction of nonreward.

Key words: dopamine; reward; attention; prediction error; inhibition; learning theory

Introduction

Using the predictive structure of the environment requires animals to learn not only about stimuli that signal the occurrence of biologically important events, such as rewards, but also about stimuli that predict the omission of expected events. Pavlovian or classical conditioning (Pavlov, 1927) provides a set of behavioral paradigms for investigating such predictive learning. Animals acquire behavioral responses to stimuli that predict rewards. Animals learn also to inhibit the responses when an additional, second stimulus occurs that predicts explicitly the omission of the reward that would have occurred otherwise (Rescorla, 1969). Thus the additional stimulus is "responsible" for the omission of reward and inhibits the behavioral response that would normally occur; it is therefore called a conditioned inhibitor.

Rewards and reward-predicting stimuli are positive reinforcers that induce approach behavior and learning, but they also have attentional functions and elicit orienting reactions that may disrupt ongoing behavior. In addition to discriminating between reward presence and absence, the conditioned inhibition paradigm sets reward and attention in direct opposition. Animals pay attention to stimuli that explicitly predict the omission, rather

than the presence, of reward and thus inhibit the response that normally occurs to the reward-predicting stimulus (Wagner and Rescorla, 1972; Rescorla, 1973). Thus reward-predicting stimuli and conditioned inhibitors have opposite reinforcing functions but commonly induce attention.

Dopamine systems play an important role in reward prediction, approach behavior, reward-directed learning, and drug addiction (Fibiger and Phillips, 1986; Wise and Hoffman, 1992; Robinson and Berridge, 1993; Robbins and Everitt, 1996), but they might also be involved in some attentional functions. Previous studies suggest that dopamine depletions impair the internal allocation of attentional resources (Brown and Marsden, 1988; Yamaguchi and Kobayashi, 1998; Granon et al., 2000), although other studies have reported conflicting results (Carli et al., 1985; Ward and Brown, 1996; Briand et al., 2001). Neurophysiological experiments on dopamine neurons have revealed reward-related, phasic activations occurring with latencies of ~100 msec after delivery of liquid or food rewards. After repeated pairing, the activation transfers to a conditioned stimulus predicting such a reward (Ljungberg et al., 1992; Hollerman and Schultz, 1998; Schultz, 1998; Waelti et al., 2001). Similar activations, which are followed usually by depressions, occur in response to attention-inducing novel or intense stimuli (Ljungberg et al., 1992; Horvitz et al., 1997). This finding might suggest a dual involvement of phasic dopamine responses in reward and attention (Schultz, 1992, 1998; Redgrave et al., 1999; Horvitz, 2000; Kakade and Dayan, 2002).

Our previous studies have investigated the coding of reward prediction in formal paradigms involving prediction errors (Waelti et al., 2001). In the present experiments we used a con-

Table 1. Experimental design

Stimuli	Pretraining	Compound conditioning	Test
Experimental	A+	AX−	X−
Control	B− (B*−, B**−)	BY−	Y−

Letters denote visual stimuli presented on a computer monitor: + denotes reward, − denotes no reward (B*− and B**− were used only during the pretraining phase, alternating semi-randomly with A+ and B−). During compound conditioning, A+, B−, AX−, and BY− alternated semi-randomly; during the test phase, all six trial types alternated semi-randomly.

ditioned inhibition paradigm to investigate whether dopamine neurons differentiate between the prediction of reward and nonreward, whether they distinguish reward from attention, and whether the prediction error coding extends to predictions of nonreward.

Materials and Methods

Two adult, female *Macaca fascicularis* monkeys (animal A, 3.5 kg; animal B, 4.5 kg) were subjected to a classic (Pavlovian) conditioning procedure with discrete trials. Reward (fruit juice) was delivered by a computer-controlled liquid valve through a spout at the animal's mouth in fixed quantities of 0.1–0.2 ml. Licking at the spout served as indicator of behavioral reactions and learning and was monitored by tongue interruptions of an infrared photobeam 4 mm below the spout. There was no specific action required by the animal for having reward delivered after a stimulus. In free reward trials, animals received a drop of juice every 12–20 sec outside of any specific task. Intertrial intervals varied semirandomly between 12 and 20 sec. Animals were moderately fluid-deprived during weekdays and returned to their home cages each day after recording. Experimental protocols conformed to the Swiss Animal Protection Law and were supervised by the Fribourg Cantonal Veterinary Office.

Behavior. During initial training, the monkeys were habituated to sit relaxed in a primate chair for several hours on each experimental day. Visual stimuli were presented for 1.5 sec on a 13 inch computer monitor at 50 cm from the monkey's eyes. Each stimulus covered a rectangle of 2.5 × 3.5°. Five artificial, complex visual stimuli, denoted as A, B, X, Y, and C, were used. We used a second set of five different stimuli to control for the contribution of sensory rather than reward stimulus components. The stimuli of the second set were also labeled as A, B, X, Y, and C, but they were presented at different locations than in the first stimulus set. Each neuron was tested with one of the two stimulus sets. To minimize generalizing reactions, each stimulus had a distinctively different and constant form, color combination, and position on the screen. Stimuli A and C were rewarded at the end of the 1.5 sec stimulus duration, whereas the other stimuli were not rewarded unless noted otherwise.

The principal training and testing in the conditioned inhibition procedure comprised three consecutive phases. In the first, pretraining phase (Table 1), stimulus A was followed by liquid reward, marked with "+" (A+), whereas stimulus B (B−) and two other stimuli (B*− and B**−) were not rewarded, denoted as "−." Stimuli B*− and B**− served to minimize generalization and were not used in subsequent stages of the task or during neuronal recordings. The four stimuli A+, B−, B*−, and B**− did not overlap in position and were presented in equal numbers of trials and in semirandom order, with maximally three consecutive trials using the same stimulus. The second, compound conditioning phase started only after neuronal recordings began. In this phase, stimulus X− was shown simultaneously with the established reward predicting stimulus A+, and no reward followed this compound (AX−). Thus stimulus X− functioned as a conditioned inhibitor that predicted the omission of the reward that was otherwise predicted by A. As a control, stimulus B− was compounded with stimulus Y− and not rewarded (BY−). A+ and B− trials continued to be used to maintain the previously established associations with reward (A+) and no reward (B−). A+, B−, AX−, and BY− trials alternated semirandomly in equal numbers of trials. During the third phase, stimuli X− and Y− were tested alone in occasional, unrewarded trials that were semirandomly intermixed at the low rate of 1:5 with A+, B−, AX−, and BY− trials to minimize learning about stimuli X and Y on these test trials. Thus, stim-

uli were not presented in blocks. Animals advanced from one phase to another once 20–25 neurons had been recorded in each phase.

Two additional control procedures served to assess the differential behavioral inhibition exerted by the fully established stimuli X− and Y− outside of the AX− and BY− stimulus compounds. In the first test, we asked whether the inhibitory power of stimulus X− would transfer to an excitatory stimulus other than the stimulus (A+) with which X− had been trained. A fully pretrained reward-predicting visual stimulus, C+, was presented together with X− or Y− in unrewarded compound trials, and anticipatory licking responses and neuronal responses were tested for differential inhibition after the CX− but not CY− stimulus compounds (summation test) (Rescorla, 1969). In a second test, we asked whether the inhibitory properties of stimulus X− were independent of being presented in compound with any other excitatory visual stimulus. Thus, at the end of the experiments, stimuli X− and Y− were paired with reward to test for differential acquisition speeds of anticipatory licking responses (learning retardation test) (Rescorla, 1969) and neuronal coding of different magnitudes of prediction errors.

Data acquisition. After completion of chair habituation training, animals were implanted under deep pentobarbital sodium anesthesia and aseptic conditions with two horizontal cylinders for head fixation and a stainless steel chamber permitting vertical, bilateral access with microelectrodes to regions of dopamine neurons. The implant was fixed to the skull with stainless steel screws, T-shaped titanium inlays, and several layers of dental cement. Animals received postoperative analgesics and antibiotics.

Using conventional electrophysiological techniques, electrical activity of single dopamine neurons was recorded extracellularly during a period of 20–60 min with moveable, glass-insulated, platinum-plated tungsten microelectrodes positioned inside a metal guide cannula. Discharges from neuronal perikarya were converted into standard digital pulses by means of an adjustable Schmitt trigger. Neurons discharged polyphasic, initially negative, or positive impulses with relatively long durations (1.8–5.5 msec) and low frequencies (0.5–8.5 impulses per second). Impulses contrasted with those of substantia nigra pars reticulata neurons (70–90 impulses per second and <1.1 msec duration), a few unknown neurons discharging impulses of <1.0 msec at low rates, and neighboring fibers (<0.4 msec duration). Recording sites in A8, A9, and A10 were determined histologically after perfusion of the animals. Only dopamine neurons activated by the presentation of the reward-predicting stimulus A+, as assessed by computer raster displays during the experiment, were used for the present experiments (~55–70% of the dopamine neuronal population). Every neuron tested with conditioned stimuli was also tested with free reward in a separate block of trials. Each neuron was recorded for a total of ~15 trials in each trial type.

Data analysis. Task-related neuronal changes were assessed within standard time windows (Ljungberg et al., 1992; Hollerman and Schultz, 1998; Waelti et al., 2001). To determine the windows, onset and offset times of activation and depression were measured for each neuron showing a change in computer raster displays. Onset time was defined by the first of three or more consecutive histogram bins in which activity deviated from baseline activity during 800 msec preceding the first trial event (bin size, 10 msec). Offset time was defined by the first of three or more consecutive bins with activity at the baseline level. Significance of change against baseline activity before the first trial event was assessed in each neuron by a matched-pairs, single-trial Wilcoxon test between onset and offset of change in neuronal activity ($p < 0.01$). Then, a common, standard time window that included 80% of onset and offset times of statistically significant changes was defined for conditioned stimuli and free reward. Standard time windows were 70–220 msec for activations and 230–570 msec for depressions after the conditioned stimuli, and 90–220 msec for activations and 60–360 msec for depressions after the time of reward.

Magnitudes of change were expressed in percentage above baseline activity preceding the first trial event in each neuron, regardless of a response being shown, by comparing the number of impulses between the standard time window and the baseline period. Conventional Wilcoxon, Mann–Whitney, and Kruskal–Wallis tests served to compare median magnitudes of change between different stimuli and situations in all
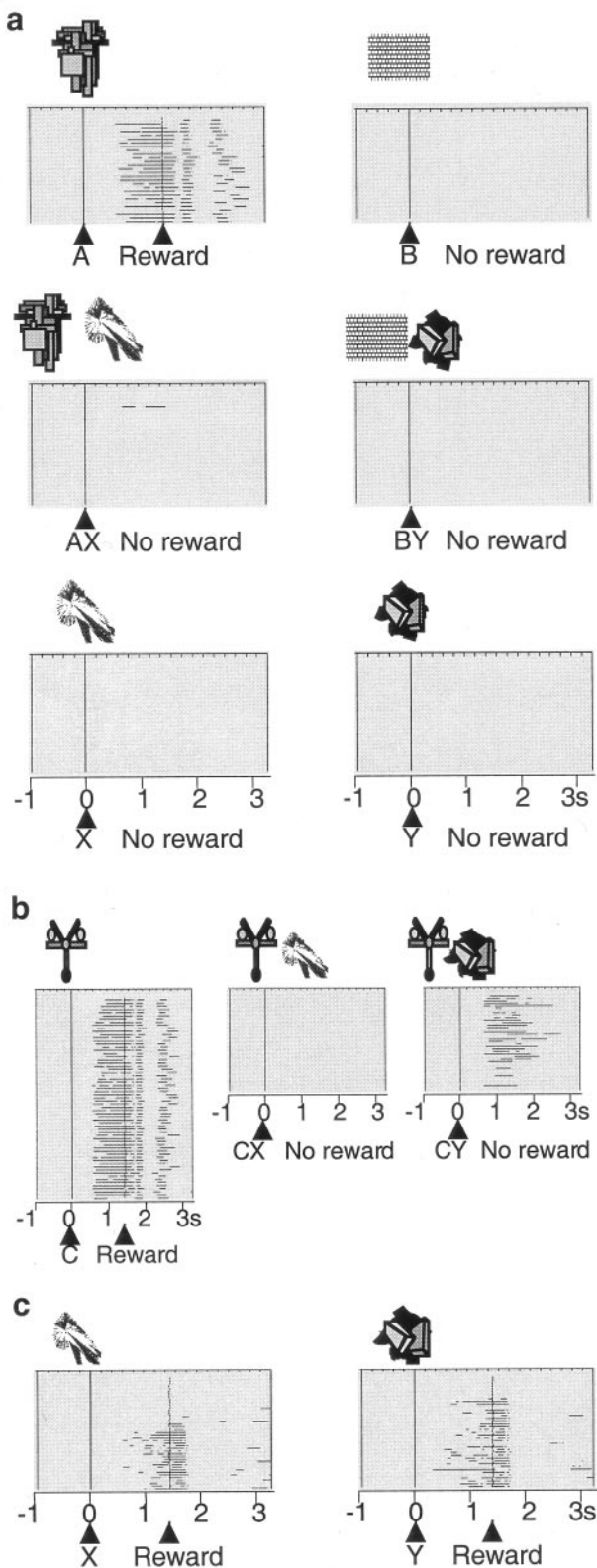
Figure 1. Behavior in the Pavlovian conditioned inhibition paradigm. *a*, Licking in the six standard trial types after learning: A+, B− pretrained stimuli (licking before and after reward occurrence to A but not to B); AX−, BY− compound stimuli (no licking); X−, Y− test (no licking). Horizontal lines indicate periods of licking in each trial; consecutive trials are shown from top to bottom. All six trial types alternated semirandomly and were separated for display. In this example, average anticipatory licking duration per A+ trial was 596 msec, similar to the overall average of 534 msec shown in Figure 2*a*. *b*, Summation test, consisting of presenting established inhibitor X− in compound with separately trained reward predictor C+.

neurons tested. Multiple Wilcoxon tests with Bonferroni corrections served for *post hoc* analysis after the Kruskal–Wallis test. The use of medians and nonparametric tests appeared appropriate, because the data were visibly not symmetrically distributed and depression magnitudes were limited. Occasional data analysis using parametric measures and tests produced essentially the same results as the used nonparametric tests. The Wilcoxon test served to determine the numbers of activated and depressed neurons by comparing activity between the standard time windows and the baseline periods ($p < 0.01$). Spearman's rank correlation coefficient, corrected for ties, served to compare the distributions of neuronal activations and depressions across midbrain groups A8–A10.

We measured anticipatory lick durations during the 1.5 sec period between the onset and offset of conditioned stimuli and lick latencies from onset of conditioned stimuli to onset of licking. Wilcoxon, Mann–Whitney, and Kruskal–Wallis tests served to compare durations and latencies between trial types and situations.

Recording sites of dopamine neurons in cell groups A8, A9, and A10 were marked with small electrolytic lesions toward the end of the experiments and reconstructed from 40-μm-thick, tyrosine hydroxylase-immunoreacted or cresyl violet-stained, stereotaxically oriented coronal sections of paraformaldehyde-perfused brains.

Results
Behavior
The Pavlovian conditioned inhibition paradigm used five different kinds of visual stimuli. Stimulus A+ predicted a drop of liquid reward, and B− served as unrewarded control stimulus. Stimulus X− was presented together with A+ and predicted, in the AX− compound, the omission of reward normally following A+ (conditioned inhibition), and Y− served as unrewarded control stimulus in compound with B− (BY−). Stimulus C+ predicted liquid reward and served specific control functions.

Monkeys licked a spout when stimulus A+ predicted a drop of liquid but not when a different stimulus (B−) predicted nothing. Median lick durations were 571 msec to stimulus A+ and 0 msec to stimulus B− (133 trial blocks; $p < 0.0001$; Wilcoxon test) (Fig. 1*a*, top). We then added stimulus X− to stimulus A+ on some trials and omitted the reward (Fig. 1*a*, middle). Licking declined in compound stimulus (AX−) trials to a median of 0 msec (89 blocks; $p < 0.0001$ compared with A+; Wilcoxon test) (Fig. 2, left) and showed longer latencies in AX− than A+ trials (medians of 937 vs 756 msec in 156 and 1306 AX− and A+ trials, respectively; $p < 0.001$; Mann–Whitney test). These data indicate that stimulus X− had become a conditioned inhibitor of the licking elicited by stimulus A+ because it appeared to be responsible for the omission of reward. Little licking occurred in unrewarded compound BY− control trials when stimulus Y− was added to stimulus B− (median, 0 msec). The occasional presentation of stimuli X− and Y− alone did not elicit licking (medians of 0 msec; 69 trial blocks) (Fig. 1*a*, bottom). The licking after the AX− and CX− stimulus compounds was as low as with stimuli that had never been rewarded (BY−). These data suggest that the inhibition produced by stimulus X− was as strong as the reward prediction generated by stimulus A+, indicating that the monkeys attended as much to the inhibitor X− as to the reward-predicting stimulus A+.

We then used two control procedures in selected test trials (see Materials and Methods) to investigate whether stimulus X− had acquired inhibitory properties independent of stimulus A+ with

←

Note absence of licking in CX− trials (behavioral inhibition). Occasional CX− and CY− test trials were interspersed with C+ trials. *c*, Learning retardation test. Acquisition of licking was slower in rewarded X+ trials than in rewarded Y+ trials. The figure shows examples from different animals. The number of licking phases after the reward varied between animals.
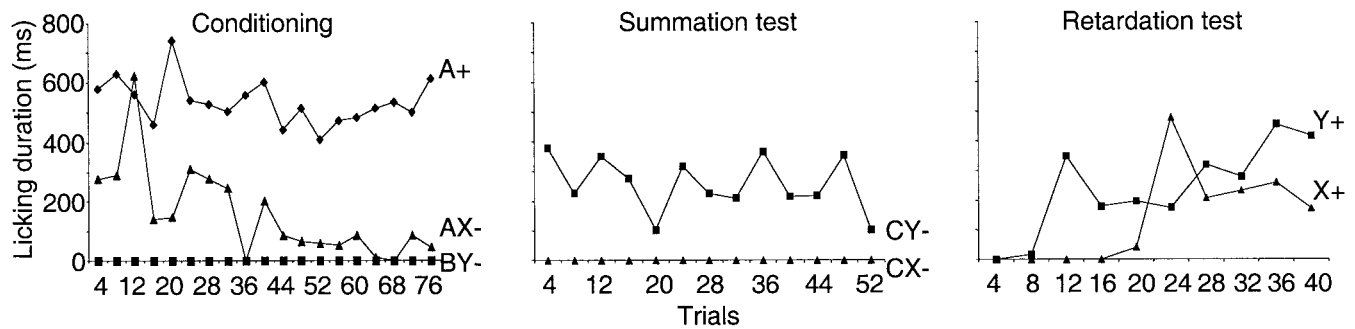
Figure 2. Acquisition of conditioned inhibition. Left, A+, AX−, and BY− trials. No licking occurred in B− trials (data not shown). Middle, CX− and CY− trials. Right, X+ and Y+ trials. The learning curves show licking during 1.5 sec stimulus duration. The unsmoothed data were averaged over four trials in each of the two animals.

which it had been trained. In the summation test we presented the established inhibitory and control stimuli X− and Y− together with established reward-predicting stimulus C+ and found that animals displayed significantly less licking in CX− than CY− or C+ trials (medians, CX−, 0 msec vs CY−, 73 msec; 120 trials; $p < 0.0001$; Wilcoxon test) (Figs. 1b, 2, middle), thus confirming the inhibitory power of stimulus X in association with another reward-predicting stimulus. In the subsequent learning retardation test we investigated whether stimulus X− had inhibitory properties outside of stimulus compounds by measuring learning speed. Pairing stimuli X and Y with reward revealed that predictive licking to stimulus X was learned slower than to stimulus Y (Figs. 1c, 2, right) and licking showed longer latencies in X+ than in Y+ trials (medians, X+, 1089 msec vs Y+, 878 msec; 35 trials; $p < 0.03$; Mann–Whitney test). Thus, rather than predicting nothing (stimulus Y−), stimulus X− had apparently acquired the capability of predicting the omission of reward on its own and without requiring a stimulus compound.

**Neural coding of conditioned inhibitors**
We tested 112 electrophysiologically and histologically characterized dopamine neurons that showed phasic activations to the pretrained stimulus A+ (17, 68, and 27 neurons in ventral midbrain groups A8, A9, and A10, respectively). With one exception, these neurons failed to respond to the unrewarded control stimulus B−, thus demonstrating neuronal discrimination between reward-predicting and neutral stimuli (Fig. 3a,c, top). Of 89 neurons tested with the unrewarded compounds AX− and BY−, 43 neurons showed a biphasic response with initial significant activation followed by significant depression to the AX− stimulus compound (Fig. 3a,c, middle), 29 showed pure significant activation, and 9 showed pure significant depression (all $p < 0.01$; Wilcoxon test). The biphasic AX− responses remained present throughout the entire experiment. The opposite pattern, depressions followed by activations, was never observed. The control stimulus compound BY− produced very few responses (one significantly activated, three significantly depressed neurons).

We tested 69 of the 89 A+-activated dopamine neurons with stimuli X− and Y− alone (10, 44, and 15 neurons in groups A8, A9, and A10, respectively) (Fig. 4), semirandomly intermixed with A+, B−, AX−, and BY− trials. Stimulus X− induced the following significant responses: pure depression (31 neurons) (Fig. 3a, bottom), weak activation followed by depression (17 neurons) (Fig. 3b, right), or weak activation alone (12 neurons). Depression followed by activation was not observed. Control stimulus Y− activated only one neuron and depressed two.

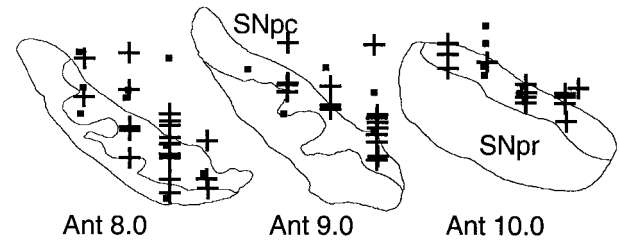Comparisons of response magnitudes in all 69 neurons

showed that depressions followed the order AX−∼X− > Y−∼B−∼BY− > A+ ($p < 0.0001$; Kruskal–Wallis test followed by Bonferroni-corrected Wilcoxon tests). Magnitudes of AX− depressions were correlated with recording locations in groups A8, A9, and A10 [Spearman's rank correlation coefficient (r) of 0.27; $p < 0.04$; stronger depressions in A8]. Depressions in X− and Y− trials were not correlated with cell groups ($r = 0.02$, $p > 0.8$; $r = -0.09$, $p > 0.4$, respectively). Furthermore, AX− depressions tended to be stronger in more laterally located neurons ($r = 0.30$; $p < 0.10$). Importantly, the inhibitory stimulus X− and compound AX− depressed dopamine activity more than control stimulus Y− (medians, X−, 35% vs Y−, 9% below baseline; $p < 0.0001$) and the control compound BY− (medians, AX−, 44% vs BY−, 5% below baseline; $p < 0.0001$). Activation magnitudes followed the order A+ > AX− > X− ($p < 0.0001$) and varied insignificantly between B−, BY−, and Y− (Fig. 3c). Magnitudes of activations were generally not correlated with recording locations in groups A8, A9, and A10 ($-0.1 < r < 0.08$; all $p > 0.4$); however, there were stronger activations in more dorsal neurons in A+ and AX− ($r = 0.33$, $p < 0.009$, for both stimuli), and stronger activations in more medial neurons in A+ and AX− trials ($r = 0.41$, $p < 0.002$; $r = 0.33$, $p < 0.01$, respectively). Activation magnitudes for stimulus X− were less than one-third of those for stimulus A+ (medians, X−, 69% in 29 neurons vs A+, 242% above baseline in 69 neurons; $p < 0.0001$). Activations to compound AX− were comparable with those elicited by stimulus A+ during initial AX− training but diminished subsequently. Thus dopamine neurons showed substantial depressions and only minor activations to conditioned inhibitors that predicted the omission of reward.

Presenting stimulus X− in compound with stimulus A+ during 1.5 sec may produce an association between the two stimuli, which would have made X− an indirect predictor of reward through the association of stimulus A+ with reward (higher-order conditioning of X with A and A with reward) (Rescorla, 1973). The indirect reward association might explain the moderate activations to stimulus X−. To test this possibility, we reduced the reward association of X− by presenting stimulus A without reward (Rescorla, 1982). This procedure gradually extinguished the behavioral and neuronal responses to stimulus A (Fig. 5a) and, after complete extinction to stimulus A, diminished the activations to AX− and X− to insignificant levels ($p > 0.25$; Mann–Whitney test). In the population of 23 neurons tested after extinction, activation to X− was reduced from 69 to 6% (Fig. 5b). Only 1 of the 23 neurons was significantly activated by stimulus X−, 2 by AX−, and 3 by A−. Free reward responses were left unchanged by this procedure (Fig. 5b, bottom), suggesting spec-
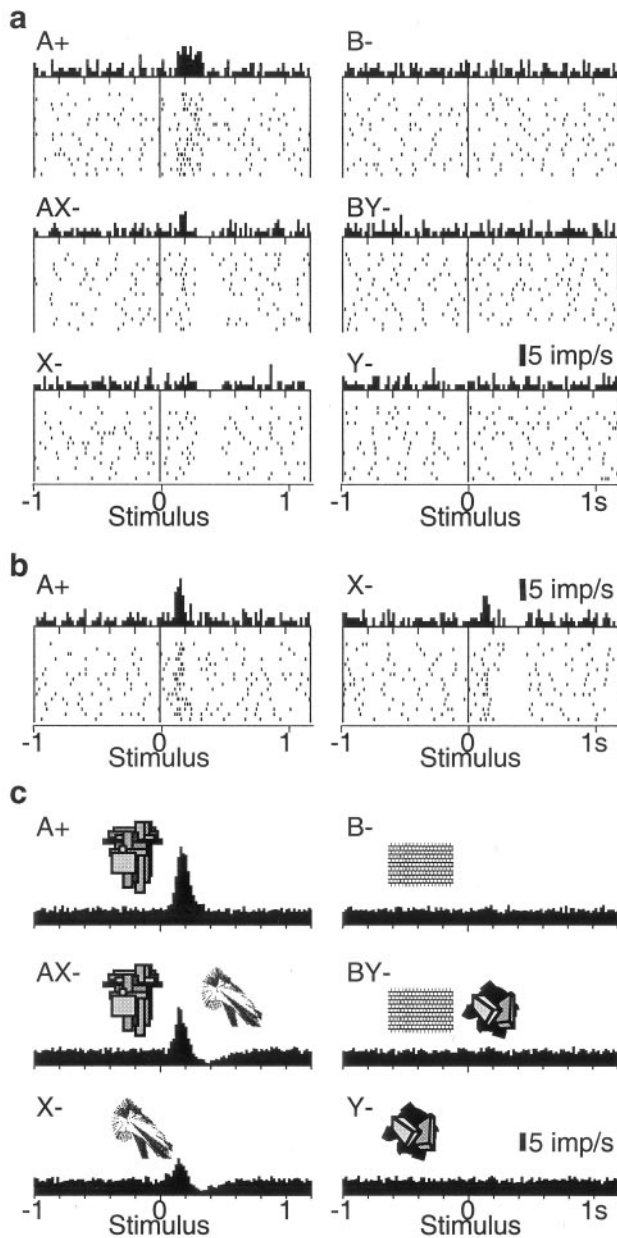
Figure 3. Responses of dopamine neurons in the conditioned inhibition paradigm. *a*, Response of one dopamine neuron in the six trial types. Note depressions after AX− and X−, and lack of activation by conditioned inhibitor X−. *b*, A case of neuronal activation by stimulus X−, which is smaller compared with A and followed by depression. *c*, Averaged population histograms of all 69 neurons tested with stimulus X−. In *a* and *b*, dots denote neuronal impulses, referenced in time to onset of stimuli. Each line of dots shows one trial, the original sequence being from top to bottom in each panel. Histograms correspond to sums of raster dots. Bin width = 10 msec. For *c*, histograms from each neuron normalized for trial number were added, and the resulting sum was divided by the number of neurons. In *a–c*, all six trial types alternated semirandomly and were separated for display.

Figure 4. Positions of dopamine neurons tested with the conditioned inhibitor (+ denotes significantly depressed neurons). Cells from two animals are superimposed on coronal sections of one animal (reconstructions from cresyl violet-stained sections). SNpc, Substantia nigra pars compacta; SNpr, substantia nigra pars reticulata; Ant 8.0 –10.0, levels anterior to the interaural stereotaxic line.

stimulus X− was eliminated. Thus the weak neuronal activations to X− before extinction of stimulus A+ appear to reflect inadvertent reward association via stimulus A+ rather than attentional processing.
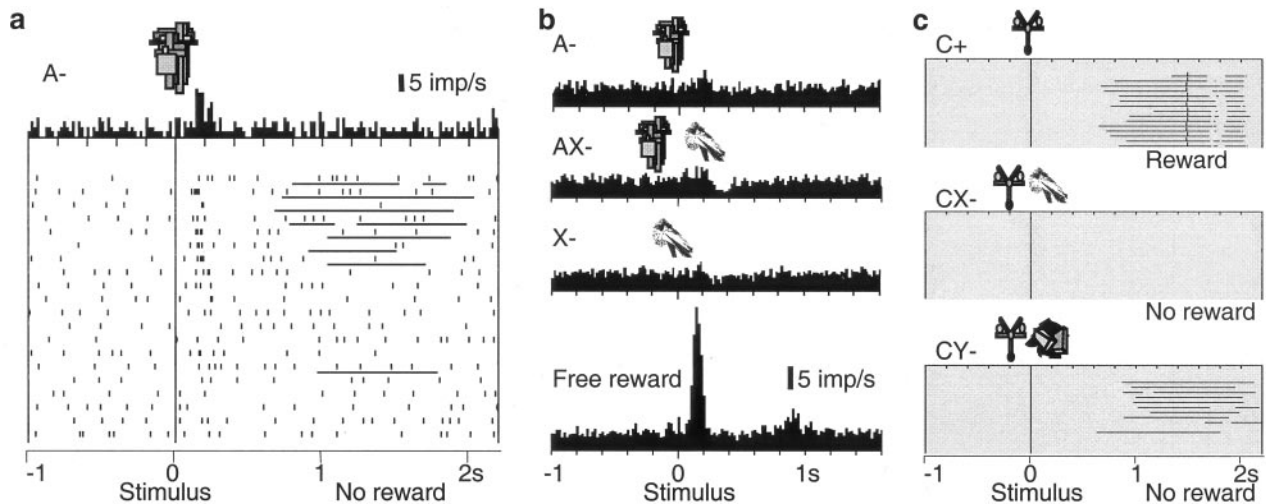
Neural coding of prediction errors

Dopamine neurons appear to code the discrepancy between the actually occurring reward and the expected reward (reward prediction error) (Waelti et al., 2001) (for review, see Schultz, 1998). We assessed whether the coding of prediction errors applied also to the prediction of reward omission, using the conditioned inhibition paradigm. This involved testing a relatively small number of neurons in specific, behaviorally, and neuronally transient situations, and these data are therefore not presented as population histograms.

The omission of reward in initial, unrewarded AX− trials of compound training constituted a negative reward prediction error, because stimulus A+ was an established reward predictor and stimulus X− was unknown. During this brief learning period, dopamine neurons showed a tendency for depression at the habitual time of reward that appeared in initial AX− trials and disappeared later (Fig. 6a, right, top vs bottom). The depression reached significance in two of the first four neurons tested in initial AX− trials in monkey A and in three of the first six neurons tested in monkey B. Dopamine activity remained uninfluenced in all BY− trials.

A depression occurred also when stimulus Y was presented in compound with an established reward-predicting stimulus C+ (summation test). Reward omission after the usually rewarded stimulus C+ constituted a negative prediction error in this situation. Dopamine neurons showed a tendency for depression at the usual reward time in initial unrewarded CY−, but not CX−, trials (Fig. 6b, right), just as in initial AX− trials (Fig. 6a, top right). The depression began to fade away with continued testing, possibly reflecting the inadvertent training of stimulus Y− as conditioned inhibitor. We tested only one neuron in each animal to avoid further training of Y.

The depression at the time of reward omission was replaced after learning by postactivation depression after compound AX−, which reached significance in 5 of 10 neurons tested in later AX− trials (Fig. 6a, bottom). Similarly, in CX− trials, a depression occurred at the time of the conditioned stimuli but not at the time of the reward (Fig. 6b, left). Although these data are limited because of the brief duration of the learning period, they may suggest that the depressions had transferred from the habitual time of the reward to the stimulus predicting the reward omission, analogous to the transfer of the dopamine activation from

ificity of response reduction. Despite the loss of neuronal activation to stimulus X−, animals continued to show no licking in compound with an independently trained reward predictor C+ (Fig. 5c). Neuronal depressions to stimuli AX− and X− were reduced (AX−, from 44 to 30%; X−, from 35 to 24% below baseline) but still remained significant ($p < 0.0001$ relative to baseline for both AX− and X−). Thus, dopamine neurons showed a weak depression, but no phasic activation, to the conditioned inhibitor X− after the potential reward association with

Figure 5. Extinction of reward-predicting stimulus removes small dopamine activation to conditioned inhibitor. *a,* Extinction training (top to bottom). Licking and neuronal responses to stimulus A− subsided gradually. Bin width = 10 msec. *b,* Average population responses in 23 dopamine neurons to A−, AX−, and X− were abolished after extinction of A. Responses to free reward were preserved in the same neurons (bottom). Trial types with conditioned stimuli alternated semirandomly and were separated for display. Control stimuli B−, BY−, and Y− continued to not elicit any responses (data not shown). *c,* Preserved conditioned behavioral inhibition after extinction of reward-predicting stimulus, as evidenced by adding the established conditioned inhibitor X− to the established reward-predicting stimulus C+ (summation test). Animals licked in initial CY− trials, indicating reward prediction but not yet conditioned inhibition (bottom).

primary reward to reward-predicting stimuli (Ljungberg et al., 1992).

We tested the surprising delivery of reward after stimulus X, investigating only one neuron in each animal to avoid conditioning of X. Because stimulus X− explicitly predicted the omission of reward, the reward after X constituted a larger prediction error than reward after the neutral stimulus Y. The behavioral learning of stimulus X+ was delayed (retardation test). The surprising delivery of reward evoked a stronger and longer lasting activation in X+ than Y+ trials (X+ trials 132% above Y+ trials; $p < 0.002$; Wilcoxon test; $n = 40$ trials; X trials 20 msec longer than Y trials; $p < 0.02$; no change in activation onset) (Fig. 6c).

The predicted omission of reward after the well learned compound AX− would not constitute a prediction error. The lack of prediction error at reward time might be constructed by a simple addition of the negative prediction error after reward omission after reward-predicting stimulus A+ and an opposite influence of stimulus X−. Hence, presentation of the established inhibitory stimulus X− alone should induce an activation at the habitual time of reward, even in the absence of reward. Such an activation of dopamine neurons, however, was never seen when stimulus X− was presented alone, similar to the neutral stimulus Y− (Fig. 6d). This result is at variance with the notion of a simple neuronal addition of prediction errors and could suggest a prediction-modulating input to dopamine neurons.

Discussion

These data suggest that reward-predicting stimuli elicit phasic activations of dopamine neurons, whereas stimuli that specifically predict the omission of reward induce minor activations and considerable depressions (Fig. 7). The extinction experiment suggests that the minor activation to the predictor of nonreward may be attributed to higher-order conditioning in the particular paradigm used. The differential responses occurred, although the animals were likely to pay similar amounts of attention to the two classes of motivationally opposite stimuli to entirely suppress the licking behavior in the presence of the conditioned inhibitor. These results indicate that the dopamine responses to reward and reward-predicting stimuli are not caused primarily by the

attention-inducing functions of rewards. The results are compatible with basic concepts of learning theory and suggest that the coding of prediction errors of dopamine neurons with reward-predicting stimuli may extend to the prediction of nonreward.

The present results suggest that the depression at the time of the unexpectedly omitted reward is transferred during inhibitory learning to the stimulus predicting the reward omission. Composite activation–depression responses occur after novel or intense stimuli (Ljungberg et al., 1992) and after stimuli resembling reward-predicting stimuli (Waelti et al., 2001). The initial, small, generalized activation of dopamine neurons to novel, intense, or reward-resembling stimuli may represent a default reward prediction that is followed immediately by an activation-canceling depression when the nonrewarding character of the stimulus is detected. A rapid detection and possible cancellation may help to signal a potentially rewarding situation until its character has been confirmed or rejected and thus could serve the function of a "motivational bonus" (Kakade and Dayan, 2002).

Attention allows differential processing of simultaneous sources of information by increasing the discriminability of stimuli or parts of stimuli (for review, see Johnston and Dark, 1986; Egeth and Yantis, 1997). Attention can be on the basis of previously acquired representations and lead to focusing on particular known events, objects, or spatial locations (Goldberg and Bruce, 1985; Moran and Desimone, 1985; Treue and Maunsell, 1996). In a different form, attention can be induced by conspicuous, physically or motivationally salient, external stimuli that elicit orienting reactions. For example, reward-predicting stimuli will engage attention. Conditioned inhibitors are also likely to induce attention, because they are capable of suppressing an otherwise occurring conditioned response. In our experiments, the conditioned inhibitor induced attention by strongly reducing licking during the conditioning phase (AX− trials) and the summation test (CX− trials).

Rewards and reward-predicting stimuli have positive reinforcing and attention-inducing functions, and dopamine neurons respond with substantial activation (Ljungberg et al., 1992). Novel and intense stimuli also have attention-inducing functions
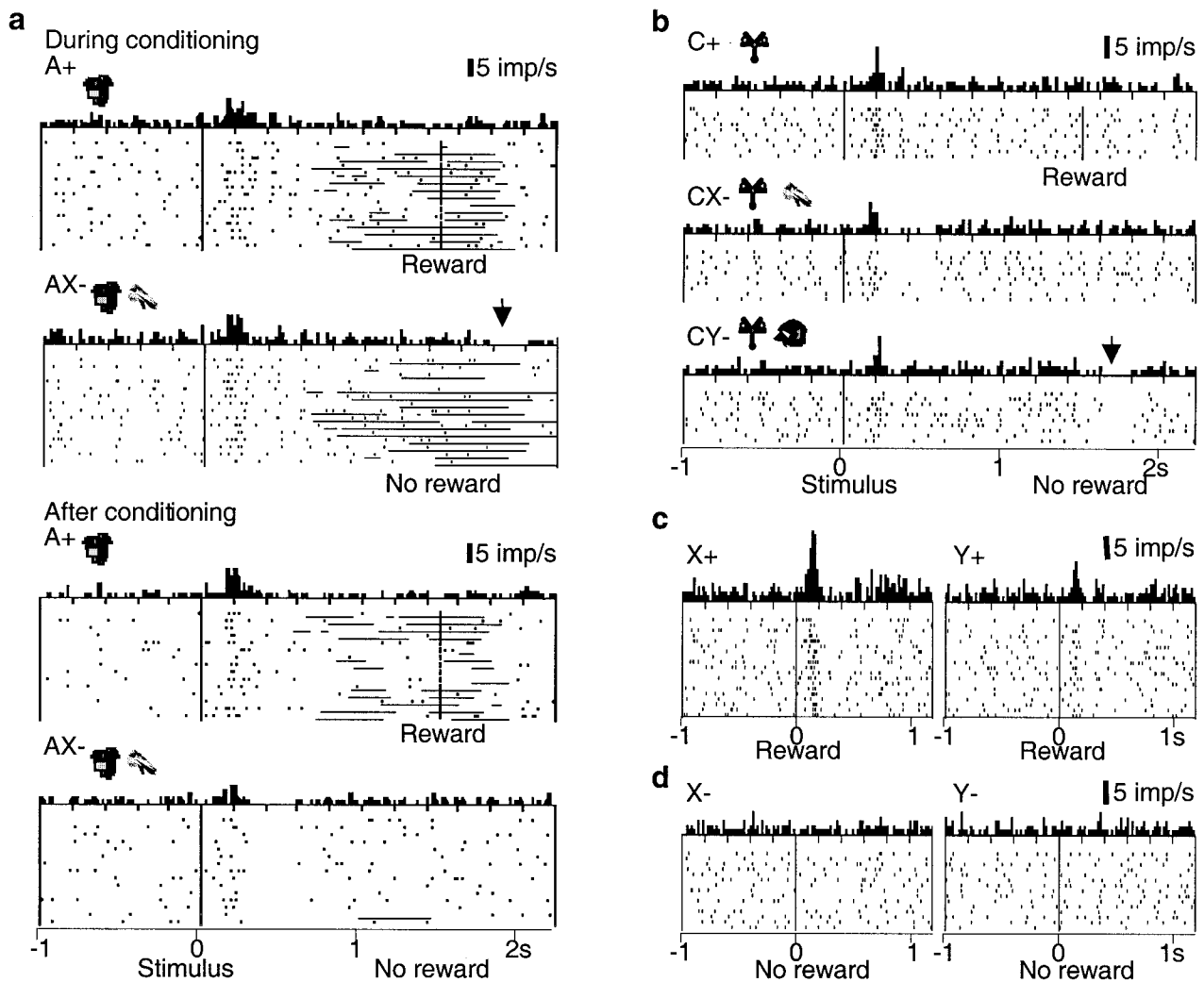
Figure 6. Dopamine coding of prediction errors in conditioned inhibition paradigm. *a*, Transfer of dopamine depression during inhibitory conditioning in AX− trials from the habitual reward time (arrow) to the conditioned compound stimulus AX− (bottom). The activation to AX− was lower compared with A+. Horizontal lines in rasters indicate periods of licking. *b*, Summation test reveals that established inhibitor X− functions independently from its original partner of compound training. A pretrained reward-predicting stimulus C+ activated dopamine neurons (top). When X− was added to this pretrained reward predictor, dopamine neurons were depressed at the time of conditioned stimuli but not at the time of expected nonreward (middle). In initial CY− control trials (bottom), dopamine neurons were depressed at time of omitted reward (arrow). *c*, Surprising reward delivery in a learning retardation test reveals neural coding of stronger reward prediction error after conditioned inhibitor X compared with neutral stimulus Y. *d*, Absence of activation of dopamine neurons at usual time of reward in trials with inhibitory stimulus X− and neutral stimulus Y−. Bin width = 10 msec.

as well as motivating and rewarding functions (Eisenberger, 1972; Humphrey, 1972; Mishkin and Delacour, 1975; Washburn et al., 1991), and dopamine neurons respond with activation followed frequently by depression (Ljungberg et al., 1992; Horvitz et al., 1997). Conditioned and primary aversive stimuli lack positive reinforcing functions but induce attention, and dopamine neurons respond only rarely with the typical phasic activations (Mirenowicz and Schultz, 1996), although they show depressions or occasional activations at a slower time scale (Schultz and Romo, 1987). Conditioned inhibitors lack positive reinforcing functions but induce attention, and dopamine neurons respond with a sequence of mild activation followed by considerable depression or, when higher-order conditioning is ruled out, only with mild depression of activity (this experiment). These results suggest that substantial, phasic activations of dopamine neurons do not occur in response to all classes of attention-inducing stimuli, counter to previous suggestions (Redgrave et al., 1999; Horvitz, 2000). A general attentional component may be reflected in the presently observed minor activations to conditioned inhibi-

tors, although the extinction experiment suggests an underlying artifact from our particular paradigm. Thus the activations to reward-predicting stimuli appear to reflect predominantly their reinforcing rather than attentional functions (Fig. 7). There might be very special forms of induced attention closely tied to reward-predicting stimuli and not occurring with any other events, although these forms would be difficult to disentangle, and it may be that more sustained dopamine activations serve some attentional functions by coding reward uncertainty (Fiorillo et al., 2003).

The present results add further evidence to the suggestion that phasic dopamine responses code a reward prediction error of the type deployed by contemporary theories of associative learning (Waelti et al., 2001). Our findings appear to extend the coding of reward prediction error to conditioned inhibition in a formal procedure. The unexpected reward after the conditioned inhibitor elicited supernormal dopamine activity. Associative theory (Rescorla and Wagner, 1972) postulates that such a reward presentation generates a larger prediction error than the presentation of the same reward in the absence of a conditioned inhibitor,

Reward coding

Attention coding

Common controls
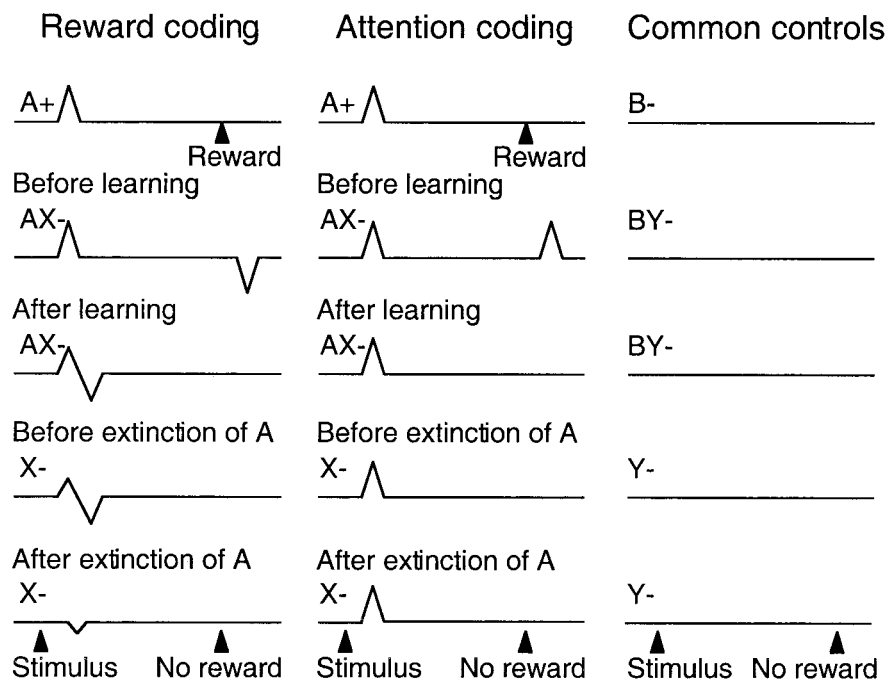
Figure 7. Schematic of observed results versus hypothetical attention coding by phasic dopamine responses in the conditioned inhibition paradigm. Stimulus A+ predicted reward and attracted attention. Compound presentation of stimuli A+ and X− led to reward omission and induced behavioral inhibition. Stimulus X− was the conditioned inhibitor that predicted reward omission and attracted attention. Thus stimuli A+ and X− had differential reward prediction but commonly attracted attention. With attentional coding they should be expected to elicit similar activating responses (middle), whereas with reward coding opposing responses were expected and observed (left). Dopamine neurons responded to stimulus A+, compound AX−, and stimulus X− in a manner compatible with reward coding (left) but not attentional coding (middle). Control stimuli B− and Y− were not associated with reward or specific attention and did not induce neuronal responses (right). Note that attentional theories of learning suggest that unidirectional but not bidirectional prediction errors drive learning through attention and would predict activations at the time of unexpected reward omission in AX− trials before learning.

and indeed it is well established that a reward presented after a conditioned inhibitor supports supernormal excitatory conditioning (Pearce and Redhead, 1995). The supernormal dopamine response might serve to recover the strong negative associative strength of the conditioned inhibitor during learning.

The prediction-error hypothesis suggests that unpredicted rewards activate dopamine cells, whereas omitted rewards depress their activity below baseline (Waelti et al., 2001). The present study shows that the depression was abolished by the presence of an established conditioned inhibitor that predicted reward omission, which also accords with the encoding of prediction errors by dopamine responses.

The observation of dopamine activation by reward-predicting stimuli is compatible with the idea that such stimuli should themselves generate a prediction error to support learning about higher-order relationships between predictive stimuli. Again, the present results extend this observation to conditioned inhibitors in that the depression induced by reward omission transferred to the stimulus that predicted this omission. This transfer of the capacity to elicit a negative prediction error encoded by the depression of dopamine activity could underlie the ability of conditioned inhibitors to support higher-order conditioning of behavioral inhibition to predictors of the conditioned inhibitor itself (Rescorla, 1976).

Despite these correspondences, one result seems at variance with simple accounts of prediction-error coding. We never found an isolated neuronal activation at habitual reward time on trials with nonrewarded presentations of the conditioned inhibitor

alone, although theoretically the absence of reward should have generated a positive prediction error (Rescorla and Wagner, 1972). The conditioned inhibitor predicted the omission of an expected reward, but on these trials, without any reward-predicting stimulus, the animal had no reason to expect reward, and a strict application of the Rescorla–Wagner theory suggests that this condition should generate a positive prediction error, which in turn should extinguish the conditioned inhibition. In this respect, however, the dopamine response corresponds to the empirical nature of behavioral inhibition rather than the theoretical analysis, in that presentations of a conditioned inhibitor alone indeed do not extinguish its inhibitory properties (Zimmer-Hart and Rescorla, 1974). It is now generally recognized that prediction-error theories of associative learning must be modified so that nonreward does not generate a prediction error in the absence of a reward expectation, a modification that is endorsed by the activity profile of the dopamine neurons.

References

Briand KA, Hening W, Poizner H, Sereno AB (2001) Automatic orienting of visuospatial attention in Parkinson's disease. Neuropsychologia 39:1240–1249.

Brown RG, Marsden CD (1988) Internal versus external cues and the control of attention in Parkinson's disease. Brain 111:323–345.

Carli M, Evenden JL, Robbins TW (1985) Depletion of unilateral striatal dopamine impairs initiation of contralateral actions and not sensory attention. Nature 313:679–682.

Egeth HE, Yantis S (1997) Visual attention: control, representation, and time course. Annu Rev Psychol 48:269–297.

Eisenberger R (1972) Explanation of rewards that do not reduce tissue needs. Psychol Bull 77:319–339.

Fibiger HC, Phillips AG (1986) Reward, motivation, cognition: psychobiology of mesotelencephalic dopamine systems. In: Handbook of physiology—the nervous system IV (Bloom FE, ed), pp 647–675. Baltimore: Williams & Wilkins.

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

Goldberg ME, Bruce CJ (1985) Cerebral cortical activity associated with the orientation of visual attention in the rhesus monkey. Vision Res 25:471–481.

Granon S, Passetti F, Thomas KL, Dalley JW, Everitt BJ, Robbins TW (2000) Enhanced and impaired attentional performance after infusion of D1 dopaminergic receptor agents into rat prefrontal cortex. J Neurosci 20:1208–1215.

Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci 1:304–309.

Horvitz JC (2000) Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. Neuroscience 96:651–656.

Horvitz JC, Stewart T, Jacobs BL (1997) Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. Brain Res 759:251–258.

Humphrey NK (1972) "Interest" and "pleasure": two determinants of a monkey's visual preferences. Perception 1:395–416.

Johnston WA, Dark VJ (1986) Selective attention. Annu Rev Psychol 37:43–75.

Kakade S, Dayan P (2002) Dopamine: generalization and bonuses. Neural Net 15:549–559.

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

Mirenowicz J, Schultz W (1996) Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. Nature 379:449–451.

Mishkin M, Delacour J (1975) An analysis of short-term visual memory in the monkey. J Exp Psychol Anim Behav Process 1:326–334.

Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. Science 229:782–784.

Pavlov IP (1927) Conditional reflexes. London: Oxford UP.

Pearce JM, Redhead ES (1995) Supernormal conditioning. J Exp Psychol Anim Behav Process 21:155–165.

Redgrave P, Prescott TJ, Gurney K (1999) Is the short-latency dopamine response too short to signal reward error? Trends Neurosci 22:146–151.

Rescorla RA (1969) Pavlovian conditioned inhibition. Psychol Bull 72:77–94.

Rescorla RA (1973) Second-order conditioning: implications for theories of learning. In: Contemporary approaches to conditioning and learning (McGuigan FJ, Lumsden DB, eds), pp 127–150. Washington, DC: Winston.

Rescorla RA (1976) Second-order conditioning of Pavlovian conditioned inhibition. Learn Motiv 7:161–172.

Rescorla RA (1982) Simultaneous second-order conditioning produces S-S learning in conditioned suppression. J Exp Psychol Anim Behav Process 8:23–32.

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

Robbins TW, Everitt BJ (1996) Neurobehavioural mechanisms of reward and motivation. Curr Opin Neurobiol 6:228–236.

Robinson TE, Berridge KC (1993) The neural basis for drug craving: an incentive-sensitization theory of addiction. Brain Res Rev 18:247–291.

Schultz W (1992) Activity of dopamine neurons in the behaving primate. Semin Neurosci 4:129–138.

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

Schultz W, Romo R (1987) Responses of nigrostriatal dopamine neurons to high intensity somatosensory stimulation in the anesthetized monkey. J Neurophysiol 57:201–217.

Treue S, Maunsell JH (1996) Attentional modulation of visual motion processing in cortical areas MT and MST. Nature 382:539–541.

Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. Nature 412:43–48.

Wagner AR, Rescorla RA (1972) Inhibition in Pavlovian conditioning: application of a theory. In: Inhibition and learning (Halliday MS, Boakes RA, eds), pp 301–336. London: Academic.

Ward NM, Brown VJ (1996) Covert orienting of attention in the rat and the role of striatal dopamine. J Neurosci 16:3082–3088.

Washburn DA, Hopkins WD, Rumbaugh DM (1991) Perceived control in rhesus monkeys (*Macaca mulatta*): enhanced video-task performance. J Exp Psychol Anim Behav Process 17:123–129.

Wise RA, Hoffman CD (1992) Localization of drug reward mechanisms by intracranial injections. Synapse 10:247–263.

Yamaguchi S, Kobayashi S (1998) Contributions of the dopaminergic system to voluntary and automatic orienting of visuospatial attention. J Neurosci 18:1869–1878.

Zimmer-Hart CI, Rescorla RA (1974) Extinction of Pavlovian conditioned inhibition. J Comp Physiol Psychol 86:837–845.