A "Good Parent" Function of Dopamine: Transient Modulation of Learning and Performance during Early Stages of Training

JON C. HORVITZ[a], **WON YUNG CHOI**[b], **CECILE MORVAN**[a], **YANIV EYNY**[b], and **PETER D. BALSAM**[c]

[a] Boston College, Department of Psychology, Chestnut Hill, Massachusetts, USA

[b] Columbia University, Department of Psychology, New York, USA

[c] Department of Psychology, Barnard College, Columbia University, and New York State Psychiatric Institute, New York, USA

Abstract

While extracellular dopamine (DA) concentrations are increased by a wide category of salient stimuli, there is evidence to suggest that DA responses to primary and conditioned rewards may be distinct from those elicited by other types of salient events. A reward-specific mode of neuronal responding would be necessary if DA acts to strengthen behavioral response tendencies under particular environmental conditions or to set current environmental inputs as goals that direct approach responses. As described in this review, DA critically mediates both the acquisition and expression of learned behaviors during early stages of training, however, during later stages, at least some forms of learned behavior become independent of (or less dependent upon) DA transmission for their expression.

Keywords

INTRODUCTION

Dopamine (DA) neurons in the substantia nigra (SN) and ventral tegmental area (VTA) respond to unexpected rewards as well as some other classes of salient events, with phasic burst-mode activity lasting several hundreds of milliseconds.[1,2] Increases in midbrain DA discharge rate lead to elevations in DA release within forebrain target sites that are disproportionally large during burst-mode firing compared to release during firing of individual action potentials.[3] Increases in DA transmission modulate both the *throughput* of corticostriatal glutamate (GLU) signals, selectively amplifying strong relative to weak GLU-mediated excitatory postsynaptic potentials,[4–6] and the *plasticity* of these synapses,[7,8] presumably corresponding to DA's roles in response expression and acquisition, respectively.[5] D1 receptor blockade reduces while D2 blockade increases striatal membrane excitability,[9] long-term potentiation (LTP),[8] and at least some forms of conditioned response acquisition.[10]

The types of learning likely to be promoted by phasic increases in DA activity depend upon whether the DA response is selective to reward, or whether DA neurons respond to salient

Address for correspondence: Jon C. Horvitz, Department of Psychology, Boston College, 301, McGuinn Hall, Chestnut Hill, MA 02467. Voice: 617-552-2999; fax: 617-552-0523., jon.horvitz@bc.edu.

events more generally. Imagine, for example, that DA-modulated plasticity increases the future likelihood that current sensory inputs receive downstream processing (beyond the level of the striatal synapse) and that motivational and behavioral relevance of the inputs is assigned at a later stage of signal processing. DA responses to the onset of a wide category of salient unexpected events could promote such a function. Similarly, such a DA response could signal to target regions that current models of response–outcome relationships must be modified (since an unexpected event occurred), without specifying the nature of the necessary modification. On the other hand, if the DA signal sets current environmental inputs as goals that direct approach responses, updates the value of goals, and/or directly increases the likelihood that the just-emitted behavior is repeated,[11–13] then DA responses would be expected to uniquely code for rewards.

Questions regarding the types of associations that underlie the acquisition and expression of operant and Pavlovian learning, and which of these are likely to be modulated by DA activity within specific forebrain target regions, have been the topic of careful analyses (see, e.g., Ref. [14]). However, for the purposes of the discussion here, we will bracket these questions, and consider a simplified model (Fig. 1) in which a GLU input (A) can evoke one of a number of outputs (B), and the DA response is necessary for an increase in the strength of the currently active input–output connection through promotion of synaptic plasticity. For the sake of this discussion, we will assume that inputs represent sensory events and output elements represent behavioral response tendencies.[15] It should be noted that because GLU inputs to striatal target sites carry information beyond simply sensory data, including information regarding expected outcomes,[5] we do not believe stimulus–response (S–R) strengthening is the sole content of DA-modulated learning. However, regardless of the content of the DA-modulated learning, if one assumes that the expression of the learned behavior involves information flow through the same set of DA-modulated GLU synapses that are strengthened as a result of learning, it becomes clear why DA's role in acquisition and expression has been so difficult to disentangle. [11,16]

As an organism seeks rewards in its environment, it will occasionally produce a response that leads to unexpected primary or conditioned reward, and undergo a strengthening of currently active synapses (Fig. 1). Under a low synaptic DA state, the likelihood of learning would be expected to be reduced for at least two reasons, corresponding to DA's role in GLU transmission and in plasticity of GLU synapses. First, reduced response output makes it less likely that the organism will encounter reward; if transmission across A→B synapses is less likely to occur, it is less likely that this connection will be strengthened. Second because DA promotes plasticity of active striatal GLU synapses, if A→B transmission were to occur and lead to reward under a low DA state, LTP would be disrupted,[8] reducing the likelihood that the active synapses (e.g., A1→B1) will undergo strengthening. The idea that reductions in DA transmission affect learning both by disrupting *synaptic plasticity* and *response expression* is of particular importance for those modeling DA's role in learning on the basis of experimental data on animal learning under conditions of pharmacologically altered DA transmission. Under such conditions, it is difficult to determine which of these two factors accounts for disruptions in response acquisition. Interestingly, one would assume that if a low synaptic DA state decreases GLU transmission across *previously strengthened* A→B connections, the expression of these connections and the corresponding learned behavior should be reduced. This appears to be true during early stages of learning, but *not* if the connection underwent extensive training prior to the reduction in DA transmission.[17] Because experimental disruptions in DA transmission affect both plasticity and response expression at least during early stages of learning, phenomena observed in slice preparations, such as opposing effects of D1 and D2 receptor antagonists on LTP, may not show correlates in behavioral learning since normal (or enhanced) learning often depends upon normal behavioral expression during the learning trials. This article will focus on (1) the conditions under which DA cells are activated, (2) the role of

D1 and D2 receptor transmission in the acquisition of a conditioned response to a sensory cue, and (3) the role of DA in the *expression* of conditioned responding during early versus late stages of training. For reasons related to experimental methodology, we will discuss DA's role in response expression prior to its role in acquisition.

REWARD VERSUS NONREWARD DA RESPONSES

The magnitude of the phasic DA response to reward is inversely related to the animal's reward *expectation* at the time that it is delivered.[2,18] Phasic responses of DA neurons can thus be said to code the discrepancy between predicted and actual reward, that is, a reward prediction error. [2,19] Like the midbrain DA response, associative learning is greatest when presentation of the unconditioned stimulus (US) is unexpected.[20,21] The phasic DA excitatory response to unexpected primary and conditioned rewards has therefore been suggested to provide a learning signal, strengthening striatal input–output connections.[2,22,23]

Like unexpected rewards, some salient nonreward events produce a phasic DA activation. VTA DA neurons show burst firing in response to nonreward auditory and visual events (light flashes and loud clicks of 1 msec duration) that are salient by virtue of their intensity and rapid onset, and for which behavioral and phasic DA responses undergo little habituation over hundreds of stimulus presentations (Fig. 2).[24] The approximately 70 msec latency to peak DA response for these stimuli is similar to those reported for unexpected primary and conditioned rewards.[2] However, the VTA DA response to salient nonreward stimuli differs from that typically observed following the presentation of primary and conditioned rewards in that the initial excitation, lasting approximately 100 msec, is typically followed by a period of inhibition lasting several hundreds of milliseconds. Salient nonreward stimuli also produce excitation followed by inhibition in DA cells of the SN.[25] Less intense stimuli that are salient by virtue of their novelty can also produce these types of excitatory–inhibitory responses.[26] In contrast, the phasic DA response to unexpected primary and conditioned rewards is characterized by an excitation for several hundreds of milliseconds, which is not followed by an inhibitory phase. [2]

Why might midbrain DA cells respond with excitation to both rewards and nonrewards in approximately 70 msec, and subsequently show an inhibition for nonrewards? Redgrave[27] noted that the 70 msec phasic DA response latency is likely to be too rapid for the reward versus nonreward status of a stimulus to be evaluated by the nervous system under natural conditions in which, for example, an animal is required to produce a saccade to foveate a visual stimulus, a response that requires longer than 70 msec. It is possible that DA cells respond to salient stimuli (e.g., stimuli that possess rapid onset, high intensity, or novelty) with an initial excitation before the reward or nonreward status of the event has been processed, and that the postexcitatory inhibition cancels out the preceding reward signal if the event is determined to be a nonreward.[28] One might speculate that the rapid DA excitatory response provides a more precise time-stamping of the event compared to that which would be possible if the DA neuron responded only after the reward/nonreward status of the event had been determined.[29]

The question of how DA cells respond to aversive events is more difficult to answer. A single-unit study found that only 11% of DA neurons respond to aversive stimuli, such as an air puff to the arm or quinine-adulterated liquid delivery[30] (but see Refs. [31] and [32]). Microdialysis studies, however, have shown that DA concentrations in forebrain target sites are increased by aversive events, such as foot shock as well as rewarding stimuli, such as food (see Refs. [1] and [33]; reviews). Some have reconciled these results by suggesting that microdialysis with its (typically) 10-min sampling period is measuring *gradual* rather than *phasic* changes in DA, and that while gradual increases in DA concentrations reflect both appetitive and aversive motivational arousal, *phasic* DA responses uniquely code a positive reward prediction error.

[2,34] Alternatively, the differing results from dialysis and single-unit studies may be reconciled by assuming that DA is increased by strong aversive stimuli (such as footshock) that have been employed in dialysis studies[35–37] and not by weak aversive stimuli (such as air puffs to the arm and quinine-adulterated saline) that have typically been employed in single-unit studies. [30] It is important to note that while DA unit responses can be shown to be phasic versus gradual, dialysis measurements may reflect either gradual changes in DA concentration or phasic increases in DA concentration, which accumulate over the course of the dialysis sampling period. It is therefore possible that the onset (or offset) of strong aversive stimuli produce phasic increases in extracellular DA concentration that accumulate over the course of the DA sampling period.

However, in a recent single-unit study that did employ strong aversive stimuli (foot pinch), midbrain DA cells did not show phasic excitatory responses, but rather showed inhibition of firing.[38] In this study, the neurochemical identity of neurons presumed to be DAergic on the basis of electrophysiological criteria was confirmed on the basis of immunopositive responses to tyrosine hydroxylase. These data lend support to the idea that phasic DA excitation is not observed for even strongly aversive events, although the fact that animals were anesthetized during the single-unit recordings calls for some caution in comparing these results to the dialysis results in which similar types of aversive stimuli increase DA release in nonanesthetized animals. The hypothesis that motivationally arousing nonreward events elicit gradual rather than phasic increases in DA activity is therefore a hypothesis that has not yet been explicitly tested. Such a test requires subsecond measurements of DA activity (e.g., fast-scan cyclic voltammetry), which can distinguish phasic from gradual changes, while exposing nonanesthetized animals to the kinds of strong aversive stimuli that have been shown to elevate DA concentrations in dialysis experiments. Nevertheless, the fact that DA cells show a phasic excitatory response to a conditioned stimulus (CS) signaling reward, a phasic excitation followed by inhibition in response to salient nonrewards, yet most DA cells do not show a phasic DA excitation to aversive stimuli (and at least in anesthetized animals show pure inhibitory responses), suggests that aversive events may fall out of the category of stimuli that generate phasic DA excitation. If so, the nonreward status of some stimuli (such as foot pinch) must be transmitted to the midbrain DA cells prior to the 70 msec onset latency typically observed for DA excitatory responses. This assumption is reasonable, for a visual stimulus that signals the absence of reward under conditions in which reward is expected (when rewarded and nonrewarded trials are intermixed) can produce DA inhibition without a preceding excitatory phase when animals are required to fixate on a central screen position prior to presentation of the visual stimulus, and when screen position itself is the property that signals reward or nonreward.[39] Thus, under conditions in which environmental constraints do not slow the processing of the nonreward status of an event, midbrain DA cells appear capable of suppressing the early excitatory response and of showing inhibition alone.

DA AND RESPONSE EXPRESSION

While the precise manner in which midbrain DA cells respond to reward and some classes of nonreward stimuli remain to be precisely characterized, it is clear that a loss of, or strong disruption in, DA transmission leads to reductions in goal-directed movement for humans[40,41] and other animals.[42–44] In Figure 1, this may be represented as a failure of B1 to become activated by A1 (or a failure of B1 to be preferentially activated compared to B outputs that are antagonist to B1, see Ref. [5]). Yet Parkinson's (PD) patients suffering nigrostriatal DA loss, show what is known as "paradoxical kinesia," that is, under certain environmental conditions, motor-impaired patients show normal or relatively normal movements. PD patients with locomotor impairment have been reported to walk quickly or even run out of a hospital room in response to a fire alarm, or to locomote normally when permitted to step over salient lines drawn on the ground.[45,46] Parkinsonian motor deficits are most pronounced when the

movement requires internal guidance, and are greatly reduced when the movement is cued by a salient external stimulus.[47,48]

In accordance with the observation that DA loss in humans leaves behavior unimpaired in the presence of strong stimulus elicitors, we observe that rats under conditions of D1 receptor blockade are strongly impaired in generating approach responses to a food compartment in the absence of a salient response-eliciting cue, but can normally initiate the same behavior when it is cued by a well-trained CS. The top panels of Figure 3 show the response of representative drug-free rats that have learned that a brief auditory CS signals the immediate delivery of a food pellet into a food compartment. The x axis of the figure represents the period 16 sec before to 10 sec after CS presentation, and each head entry occurring during this period is depicted as a horizontal line. The y axis represents successive trials (1–28) from the bottom to the top of the y axis. The bottom panels show behavior during test sessions in which the selective D1 antagonist SCH23390[49] was administered prior to the session. Figure 3A shows that after 3 days of training, the drug disrupts both cued and noncued responding. However, after 16 daily sessions of 28-trial/session (approximately 450 total) CS–food pairings, the D1 antagonist (a) continues to reduce the frequency of noncued head entries during the inter-trial interval, but (b) produces no impairment in the latency to perform the same head entry behavior in response to the CS (Fig. 3B). The DA-independent initiation of the approach response requires that the CS be well acquired, because approach responses to the CS during early stages of training are highly vulnerable to D1 receptor blockade.[17]

During this early stage of training, D1 receptor blockade does not simply produce a slowing of the approach response, but rather, appears to increase the likelihood that the animals will not emit a cue-elicited approach response. Trials on which the animal shows a latency of greater than 10 sec to emit a head entry following cue presentation are designated "missed trials." As can be seen in the bottom panel of Figure 3A, SCH-treated animals show a large number of "missed" trials (rows with no horizontal lines for 0 to 10 sec after CS presentation). Figure 4 shows that during a day 3 test session, the D1 antagonist produces a dose-dependent increase in the proportion of test trials that animals "miss." However, the D1 antagonist does not increase misses (or latencies) on days 7 or 17 of training. Because animals receive only a single drug injection, the change in response vulnerability to the D1 antagonist cannot be attributed to repeated drug administration. Further, the reduced vulnerability of the conditioned response to D1 antagonist challenge is not due to a reduced overall behavioral effectiveness of the drug in well-trained animals, for the drug suppresses locomotion (Fig. 5), as well as noncued head entries, during all phases of learning. The reduced vulnerability to D1 antagonist challenge is specific to the conditioned cue-elicited response. The cued approach is specifically mediated by D1 rather than D2 receptor activity, for during early phases of learning, the cued approach response is disrupted by D1 and not by D2 receptor antagonist treatment[50] and during later stages, it is disrupted by neither treatment.[17,51]

This raises the question of what type of change is occurring within DA target regions such that a behavioral response that was previously D1-dependent becomes independent of (or less dependent upon) DA transmission. At least two general kinds of change may be possible. First, it is possible that with extended training, the behavioral response shifts to mediation by non-DA target areas, and therefore becomes less subject to DA modulation. It has been suggested that over the course of habit learning, learned sensory-motor representations may shift from corticostriatal–basal ganglia circuits to direct corticocortical mediation.[52,53] Alternatively, it is possible that as the conditioned behavioral response becomes well acquired, its expression continues to be mediated by the same neurons that originally mediated response expression, but DA plays a declining role in modulating that expression. For example, cortical GLU input to striatal cells may depend on D1 transmission to amplify the strength of task-relevant input signals (A→B1 in Fig. 1)[5,54] during early stages of learning. During later stages of learning,

these GLU synapses may become so efficient that DA facilitation of GLU transmission is no longer necessary for normal responding.

To precisely characterize the nature of the changes that accompany the shift to DA-independent performance during *later* stages of training, it will be necessary to anatomically localize the site of DA's (D1-dependent) mediation of performance during *earlier* stages. Work in our laboratory by Won Yung Choi has been directed toward identifying the central site(s) of D1-mediated conditioned approach performance, an effort made difficult by the fact that DA forebrain target sites within which D1 receptor blockade reduces rates of lever pressing during early stages of learning (the nucleus accumbens core and prefrontal cortex[55,56]) are not substrates for D1-mediated performance of the conditioned approach.[57,58] Related to the goal of identifying the central site(s) of D1-mediated response expression during early stages of learning is the question of whether the shift to DA-independent performance occurs not only for the conditioned approach described above, but extends to other forms of learning (e.g., the operant lever press). If so, it is clear that the time course for the shift in the more complex lever press response is much longer than that for the simple cued approach.[59]

D1 PROMOTION OF APPETITIVE LEARNING

D1 and D2 receptor transmission has been shown to produce opposite effects on DA-modulated LTP. D1 antagonism reduces[7,60] while D2 antagonism or D2 receptor knockout[61,62] increases the magnitude of LTP. These opposite effects of D1 and D2 receptor blockade on LTP appear to result from their opposing effects on those intracellular cascades within striatal neurons, which lead to dopamine and cAMP-regulated phosphoprotein of 32 kDa (DARPP-32) phosphorylation. D1 receptor binding causes a G protein-mediated increase in the activity of adenylyl cyclase, increased cAMP formation, stimulation of protein kinase A, and phosphorylation (i.e., activation) of DARPP-32.[63,64] D2 binding produces the opposite effect on DARPP-32 activation, both by reducing adenylyl cyclase activity and through a mechanism that involves calcineurin activation.[64,65] It is known that the activation and deactivation state of DARPP-32 critically mediates the facilitative and inhibitory effects of D1 and D2 binding on LTP induction, respectively.[7,8] Thus, via opposing effects on the activation state of DARPP-32, D1 activity promotes and D2 activity restricts LTP.

However, the effect of D1 and D2 receptor blockade on behavioral learning is difficult to examine experimentally because D1 and D2 antagonist-induced disruptions in behavioral performance may indirectly lead to poor behavioral acquisition. If one were to administer a D1 or D2 antagonist on day 1 of cued approach training, many animals would fail to approach the food compartment at all during the session, and almost all animals would show very long latencies to respond to the CS (unpublished observations). It would be impossible to know whether these behavioral reductions were due to impaired learning or performance. If one were to test these animals during a drug-free test session on day 2, one would still be unable to determine whether increased latencies to respond to the CS on day 2 were the result of a day 1 learning deficit or the long CS–US intervals and fewer CS–US trials that the neuroleptic-treated rats experienced on day 1. (In Fig. 1, failure to express A1→B1 on day 1 precludes the strengthening of the connection.)

Even the place preference paradigm, which is one of the least problematic for assessing the effect of neuroleptic treatment on reward conditioning, suffers from the possibility of performance disruptions indirectly producing interference with learning. If DA antagonist treatment reduces exploratory behavior[17,58,66] during the sessions in which a distinctive environment is paired with reward, it will be difficult to determine whether apparent neuroleptic-induced disruptions of conditioning results from a disruption of learning or from a reduction in the animal's sampling of the CS, that is, the contextual stimuli of the environment.

One way to avoid this motor confound is to administer D1 or D2 antagonists *after* a learning session, since a disruption in learning (assessed during a later session when the drug is no longer behaviorally active) cannot be accounted for by behavioral disruptions during the learning session. Postsession D1 antagonist administration to the nucleus accumbens does disrupt some forms of appetitive learning[67] (but see Ref. [68]). However, posttraining disruptions in DA transmission would only be expected to disrupt acquisition if DA facilitation of learning is due to tonic levels of DA transmission, which promote consolidation of information after the learning trials have terminated. To the extent that DA's role in learning is due to phasic DA responses at the time that the unexpected reward is presented (see above), DA antagonist administration after the conditioning session would not be expected to affect learning.

To ask whether D1 or D2 receptor blockade produces a disruption in the acquisition of appetitive learning, we took advantage of the fact that animals receiving 17 days of CS–food pairings in a nondrugged state show no reduction in the number of food pellets retrieved and no increase in the CS–food interval even when tested under relatively high doses of D1 or D2 antagonist drugs.[51,58] We therefore exposed animals to 17 days of training, in which the sound of the feeder signaled food delivery (28 trials per day, with a variable time 70 sec-inter-trial interval). On day 18, a tone was presented 3 sec prior to (and terminated 3 sec after) food delivery for animals pretreated with vehicle, D1 or D2, antagonist drugs. Another group received unpaired presentations of tone and food. On the following tone-alone drug-free test day (day 19), we examined approach responses elicited by the tone. As can be seen in Figure 6 (top panel), the high (0.16 mg/kg) dose of D1 antagonist SCH23390, administered during the day 18 conditioning session, reduced the animals' approach responses to the tone during the day 19 drug-free test session compared to animals that were not drugged during the conditioning session. The D2 antagonist raclopride during the day 18 conditioning session did not reduce approach responses to the tone on day 19, but rather produced a dose-dependent *increase* in responses to the CS. Importantly, neither the SCH23390 nor the raclopride groups showed increased latencies to retrieve the food pellet during the day 18 conditioning session.[10]

This D1 antagonist-induced disruption, and D2 antagonist-induced facilitation, of conditioned responding mirrors the effects of D1 and D2 antagonist drugs on striatal LTP (see Ref. [8]) under conditions in which key parameters of conditioning (number of CS–US pairings and latency to respond to the CS) were themselves unaffected by the DA antagonists during the conditioning session. The finding that D1 receptor blockade reduced conditioning is consistent with a number of other studies showing that D1 receptor blockade reduces LTP[7,8] and disrupts operant and Pavlovian conditioning.[56,67,69,70] The finding that D2 receptor blockade increases conditioning mirrors the effects of D2 antagonism on LTP, suggesting that these opposite effects of D1 and D2 receptor blockade on learning may reflect their opposing effects on LTP.

However, there are two other plausible explanations for the observed D2 antagonist-induced promotion of conditioning. First, as noted above, to ensure that neuroleptic treatment did not disrupt approach responses during the conditioning session, animals received magazine training sessions for 17 days prior to the day 18 tone-food conditioning session. Conditioning of the context is strong after this sort of training.[71,72] Additionally, this procedure amounts to overtraining on CS1 (magazine) →food, and then conditioning CS2 (tone) → CS1(magazine) →food. It is possible that conditioning to the tone was to some degree attenuated by blocking, that is, because the context and the sound of the magazine already signaled food, the occurrence of the food was less surprising and less likely to strengthen the associative value of the tone than it would have been in the absence of overtraining.[73] A D2 antagonist-induced reduction in animals' *expectation* of food in the experimental context or following the sound of the magazine during the day 18 conditioning session could have reduced blocking and thereby produced increased conditioning to the tone compared to nondrugged animals. A second

alternative explanation for the D2 antagonist-induced increase in conditioning is that while the D2 antagonist blocks postsynaptic D2 receptors, it also blocks D2 autoreceptors, leading to an increase in DA release at forebrain target sites.[74,75] Under conditions of D2 receptor blockade, D1 receptors, which remain available for DA binding, may be exposed to an increased amount of DA compared to the D1 occupancy for nondrugged animals. If one assumes that D1 transmission facilitates appetitive learning, as the present SCH23390 results and other studies cited above suggest, both D1 antagonist and D2 antagonist results can be accounted for by a D1-mediated promotion of learning.

SUMMARY

We have highlighted three issues that we believe merit careful experimental examination to understand more precisely the relationship between DA neuronal firing, response acquisition, and response expression: (1) While extracellular DA concentrations are increased by a wide category of salient events,[1] there is evidence to suggest that DA responses to primary and conditioned rewards may be distinct from those elicited by other types of salient events. This reward-specific mode of responding is necessary if DA functions to strengthen response tendencies adaptively under particular environmental conditions. (2) There is recent evidence that for some forms of neural plasticity (LTP) and learning (conditioned approach), interruption of D1 receptor transmission reduces while blockade of D2 receptors promotes plasticity and learning, suggesting that the two DA receptor families may play opposing roles in at least some forms of learning. Because data on opposing D1 and D2 roles in these functions come from a relatively small and recent set of findings, we view the claim as an early-stage hypothesis that merits further experimental examination, but one of potential importance. (3) After a large number of conditioning trials, at least some behaviors become less dependent upon DA transmission for their expression. It will be important to determine whether this reduced dependence upon DA transmission reflects a strengthening of GLU synapses during learning, such that DA is no longer needed to modulate synaptic transmission, or whether, as some have suggested, behaviors that were once represented in DA-innervated regions come to be represented in anatomically distinct regions that are not subject to DA's modulatory influence.

It is known that extended training of an operant response leads to a shift from "outcome-mediation" to S–R (automatized) performance.[76] After a small number of training sessions, animals for whom the outcome (e.g., food) value is degraded (e.g., when food is paired with LiCl-induced illness) show reduced vigor of operant responding (compared to control animals that received unpaired exposure to illness and food). In these procedures, devaluation of the outcome takes place when the animal does not have the opportunity to emit the learned behavioral response; that is, the response itself is not "punished." Reductions in the rate of responding during later sessions therefore reflect a reduction in the animal's representation of the outcome value, and demonstrate that the outcome representation modulates expression of the learned behavior. After extended training, at least some learned behaviors become insensitive to outcome devaluation, pointing to a shift from outcome-mediated to automatized responding.[76] It has recently been shown that amphetamine-induced sensitization of DA receptors prior to operant lever press training causes animals to undergo an abnormally rapid shift from outcome-mediated to S–R responding,[77] suggesting that DA transmission plays a role in the shift to automatized behavior. Current studies in our laboratory by Cecile Morvan suggest that the shift from outcome-mediated to S–R responding coincides temporally with the shift to DA-independent behavioral performance. Based upon the data above, it seems plausible that DA's role in response acquisition and expression resembles that of a "good parent" who, during early stages of learning, promotes behavioral acquisition and expression, permitting the behavior to finally be performed without DA's involvement.

References

1. Horvitz JC. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. Neuroscience 2000;96:651–656. [PubMed: 10727783]

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

3. Suaud-Chagny MF, et al. Relationship between dopamine release in the rat nucleus accumbens and the discharge activity of dopaminergic neurons during local *in vivo* application of amino acids in the ventral tegmental area. Neuroscience 1992;49:63–72. [PubMed: 1357587]

4. Cepeda C, et al. Dopaminergic modulation of NMDA-induced whole cell currents in neostriatal neurons in slices: contribution of calcium conductances. J Neurophysiol 1998;79:82–94. [PubMed: 9425179]

5. Horvitz JC. Dopamine gating of glutamatergic sensorimotor and incentive smotivational input signals to the striatum. Behav Brain Res 2002;137:65–74. [PubMed: 12445716]

6. Kiyatkin EA, Rebec GV. Dopaminergic modulation of glutamate-induced excitations of neurons in the neostriatum and nucleus accumbens of awake, unrestrained rats. J Neurophysiol 1996;75:142–153. [PubMed: 8822548]

7. Calabresi P, et al. Dopamine and cAMP-regulated phosphoprotein 32 kDa controls both striatal long-term depression and long-term potentiation, opposing forms of synaptic plasticity. J Neurosci 2000;20:8443–8451. [PubMed: 11069952]

8. Centonze D, et al. Dopaminergic control of synaptic plasticity in the dorsal striatum. Eur J Neurosci 2001;13:1071–1077. [PubMed: 11285003]

9. West AR, Grace AA. Opposite influences of endogenous dopamine D1 and D2 receptor activation on activity states and electrophysiological properties of striatal neurons: studies combining *in vivo* intracellular recordings and reverse microdialysis. J Neurosci 2002;22:294–304. [PubMed: 11756513]

10. Eyny YS, Horvitz JC. Opposing roles of D1 and D2 receptors in appetitive conditioning. J Neurosci 2003;23:1584–1587. [PubMed: 12629161]

11. Ettenberg A. Dopamine, neuroleptics and reinforced behavior. Neurosci Biobehav Rev 1989;13:105–111. [PubMed: 2573021]

12. Wise R. Neuroleptics and operant behavior: the anhedonia hypothesis. Behav Brain Sci 1982;5:39–87.

13. Wise RA, Rompre PP. Brain dopamine and reward. Annu Rev Psychol 1989;40:191–225. [PubMed: 2648975]

14. Balleine BW. Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. Physiol Behav 2005;86:717–730. [PubMed: 16257019]

15. Wickens JR, Reynolds JN, Hyland BI. Neural mechanisms of reward-related motor learning. Curr Opin Neurobiol 2003;13:685–690. [PubMed: 14662369]

16. Horvitz JC, Ettenberg A. Haloperidol blocks the response-reinstating effects of food reward: a methodology for separating neuroleptic effects on reinforcement and motor processes. Pharmacol Biochem Behav 1988;31:861–865. [PubMed: 3252277]

17. Choi WY, Balsam PD, Horvitz JC. Extended habit training reduces dopamine mediation of appetitive response expression. J Neurosci 2005;25:6729–6733. [PubMed: 16033882]

18. Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. Science 2003;299:1898–1902. [PubMed: 12649484]

19. Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 1996;16:1936–1947. [PubMed: 8774460]

20. Kamin, LJ. Fundamental Issues in Associative Learning. Dalhousie University Press; Halifax, Nova Scotia: 1969. Selective Association and Conditioning.

21. Mackintosh NJ. A theory of attention: variations in the associability of stimulus with reinforcement. Psychol Rev 1975;82:276–298.

22. Beninger RJ. The role of dopamine in locomotor activity and learning. Brain Res 1983;287:173–196. [PubMed: 6357357]

23. Wickens J. Striatal dopamine in motor activation and reward-mediated learning: steps towards a unifying model. J Neural Transm Gen Sect 1990;80:9–31. [PubMed: 2407269]

24. Horvitz JC, Stewart T, Jacobs BL. Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. Brain Res 1997;759:251–258. [PubMed: 9221945]

25. Steinfels GF, et al. Response of dopaminergic neurons in cat to auditory stimuli presented across the sleep-waking cycle. Brain Res 1983;277:150–154. [PubMed: 6640288]

26. Schultz W, Romo R. Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. J Neurophysiol 1990;63:607–624. [PubMed: 2329364]

27. Redgrave P, Prescott TJ, Gurney K. Is the short-latency dopamine response too short to signal reward error? Trends Neurosci 1999;22:146–151. [PubMed: 10203849]

28. Kakade S, Dayan P. Dopamine: generalization and bonuses. Neural Netw 2002;15:549–559. [PubMed: 12371511]

29. Redgrave P, Gurney K. The short-latency dopamine signal: a role in discovering novel actions? Nat Rev Neurosci 2006;7:967–975. [PubMed: 17115078]

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

31. Guarraci FA, Kapp BS. An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential Pavlovian fear conditioning in the awake rabbit. Behav Brain Res 1999;99:169–179. [PubMed: 10512583]

32. Kiyatkin EA. Functional properties of presumed dopamine-containing and other ventral tegmental area neurons in conscious rats. Int J Neurosci 1988;42:21–43. [PubMed: 3209370]

33. Salamone JD. The involvement of nucleus accumbens dopamine in appetitive and aversive motivation. Behav Brain Res 1994;61:117–133. [PubMed: 8037860]

34. Daw ND, Kakade S, Dayan P. Opponent interactions between serotonin and dopamine. Neural Netw 2002;15:603–616. [PubMed: 12371515]

35. Abercrombie ED, et al. Differential effect of stress on *in vivo* dopamine release in striatum, nucleus accumbens, and medial frontal cortex. J Neurochem 1989;52:1655–1658. [PubMed: 2709017]

36. Sorg BA, Kalivas PW. Effects of cocaine and footshock stress on extracellular dopamine levels in the ventral striatum. Brain Res 1991;559:29–36. [PubMed: 1782559]

37. Young AM, Joseph MH, Gray JA. Latent inhibition of conditioned dopamine release in rat nucleus accumbens. Neuroscience 1993;54:5–9. [PubMed: 8515846]

38. Ungless MA, Magill PJ, Bolam JP. Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. Science 2004;303:2040–2042. [PubMed: 15044807]

39. Takikawa Y, Kawagoe R, Hikosaka O. A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward mapping. J Neurophysiol 2004;92:2520–2529. [PubMed: 15163669]

40. Marsden CD. Which motor disorder in Parkinson's disease indicates the true motor function of the basal ganglia? Ciba Found Symp 1984;107:225–241. [PubMed: 6568150]

41. Stern ER, et al. Maintenance of response readiness in patients with Parkinson's disease: evidence from a simple reaction time task. Neuropsychology 2005;19:54–65. [PubMed: 15656763]

42. Carli M, Evenden JL, Robbins TW. Depletion of unilateral striatal dopamine impairs initiation of contralateral actions and not sensory attention. Nature 1985;313:679–682. [PubMed: 3974701]

43. Fowler SC, Liou JR. Haloperidol, raclopride, and eticlopride induce microcatalepsy during operant performance in rats, but clozapine and SCH 23390 do not. Psychopharmacology (Berl) 1998;140:81–90. [PubMed: 9862406]

44. Salamone JD, Correa M. Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. Behav Brain Res 2002;137:3–25. [PubMed: 12445713]

45. Jahanshahi M. Willed action and its impairments. Cogn Neuropsychol 1998;15:483.

46. Martin, JP. The Basal Ganglia and Posture. Pitman Medical; London: 1967.

47. Frischer M. Voluntary vs autonomous control of repetitive finger tapping in a patient with Parkinson's disease. Neuropsychologia 1989;27:1261–1266. [PubMed: 2594172]

48. Schettino LF, et al. Deficits in the evolution of hand preshaping in Parkinson's disease. Neuropsychologia 2004;42:82–94. [PubMed: 14615078]

49. Iorio LC, et al. SCH 23390, a potential benzazepine antipsychotic with unique interactions on dopaminergic systems. J Pharmacol Exp Ther 1983;226:462–468. [PubMed: 6135795]

50. Choi W, Horvitz JC. D1 but not D2 Receptor mediation of the expression of a Pavlovian approach response. Society for Neuroscience Abstracts 2003;29:716.1.

51. Horvitz JC, Eyny YS. Dopamine D2 receptor blockade reduces response likelihood but does not affect latency to emit a learned sensory-motor response: implications for Parkinson's disease. Behav Neurosci 2000;114:934–939. [PubMed: 11085607]

52. Ashby FG, Ennis JM, Spiering BJ. A neurobiological theory of automaticity in perceptual categorization. Psychol Rev. 2007 In press.

53. Carelli RM, Wolske M, West MO. Loss of lever press-related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. J Neurosci 1997;17:1804–1814. [PubMed: 9030639]

54. O'Donnell P. Dopamine gating of forebrain neural ensembles. Eur J Neurosci 2003;17:429–435. [PubMed: 12581161]

55. Baldwin AE, Sadeghian K, Kelley AE. Appetitive instrumental learning requires coincident activation of NMDA and Dopamine D1 Receptors within the medial prefrontal cortex. J Neurosci 2002;22:1063–1071. [PubMed: 11826135]

56. Smith-Roe SL, Kelley AE. Coincident activation of NMDA and dopamine D1 receptors within the nucleus accumbens core is required for appetitive instrumental learning. J Neurosci 2000;20:7737–7742. [PubMed: 11027236]

57. Choi, W. Dissertation. Columbia University; New York: 2005. The role of D1 and D2 dopamine receptors in the expression of a simple appetitive response at different stages of learning.

58. Choi W, Horvitz JC. Functional dissociations produced by D1 antagonist SCH 23390 infusion to the nucleus accumbens core versus dorsal striatum under an appetitive approach paradigm. Society for Neuroscience Abstracts 2002:28.

59. Morvan CI, et al. The shift to dopamine-independent expression of an overtrained pavlovian approach response coincides with the shift to S-R performance. Society for Neuroscience Abstracts 2006:463.10.

60. Kerr JN, Wickens JR. Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum *in vitro*. J Neurophysiol 2001;85:117–124. [PubMed: 11152712]

61. Calabresi P, et al. Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. J Neurosci 1997;17:4536–4544. [PubMed: 9169514]

62. Yamamoto Y, et al. Expression of N-methyl-D-aspartate receptor-dependent long-term potentiation in the neostriatal neurons in an *in vitro* slice after ethanol withdrawal of the rat. Neuroscience 1999;91:59–68. [PubMed: 10336060]

63. Hemmings HC Jr, Nairn AC, Greengard P. DARPP-32, a dopamine-and adenosine 3′:5′-monophosphate-regulated neuronal phosphoprotein. II. Comparison of the kinetics of phosphorylation of DARPP-32 and phosphatase inhibitor 1. J Biol Chem 1984;259:14491–14497. [PubMed: 6501303]

64. Nishi A, Snyder GL, Greengard P. Bidirectional regulation of DARPP-32 phosphorylation by dopamine. J Neurosci 1997;17:8147–8155. [PubMed: 9334390]

65. Greengard P, Allen PB, Nairn AC. Beyond the dopamine receptor: the DARPP-32/protein phosphatase-1 cascade. Neuron 1999;23:435–447. [PubMed: 10433257]

66. Ahlenius S, et al. Suppression of exploratory locomotor activity and increase in dopamine turnover following the local application of cis-flupenthixol into limbic projection areas of the rat striatum. Brain Res 1987;402:131–138. [PubMed: 3828778]

67. Dalley JW, et al. Time-limited modulation of appetitive Pavlovian memory by D1 and NMDA receptors in the nucleus accumbens. Proc Natl Acad Sci USA 2005;102:6189–6194. [PubMed: 15833811]

68. Hernandez PJ, et al. AMPA/kainate, NMDA, and dopamine D1 receptor function in the nucleus accumbens core: a context-limited role in the encoding and consolidation of instrumental memory. Learn Mem 2005;12:285–295. [PubMed: 15930507]

69. Azzara AV, et al. D1 but not D2 dopamine receptor antagonism blocks the acquisition of a flavor preference conditioned by intragastric carbohydrate infusions. Pharmacol Biochem Behav 2001;68:709–720. [PubMed: 11526968]

70. Beninger RJ, Miller R. Dopamine D1-like receptors and reward-related incentive learning. Neurosci Biobehav Rev 1998;22:335–345. [PubMed: 9579323]

71. Balsam PD, Gibbon J. Formation of tone-US associations does not interfere with the formation of context-US associations in pigeons. J Exp Psychol Anim Behav Process 1988;14:401–412. [PubMed: 3183580]

72. Balsam PD, Schwartz AL. Rapid contextual conditioning in autoshaping. J Exp Psychol Anim Behav Process 1981;7:382–393. [PubMed: 7288369]

73. Kamin, LJ. Fundamental Issues in Instrumental Learning. Mackintosh, NJ.; Honig, WK., editors. Dalhousie University Press; Halifax, Nova Scotia: 1969. p. 42-64.

74. Altar CA, et al. Dopamine autoreceptors modulate the *in vivo* release of dopamine in the frontal, cingulate and entorhinal cortices. J Pharmacol Exp Ther 1987;242:115–120. [PubMed: 3612520]

75. Ungerstedt U, et al. Functional classification of different dopamine receptors. Psychopharmacology Suppl 1985;2:19–30. [PubMed: 3159010]

76. Dickinson A. Actions and habits: the development of behavioral autonomy. Phil Trans R Soc London (biol) 1985;308:67–78.

77. Nelson A, Killcross S. Amphetamine exposure enhances habit formation. J Neurosci 2006;26:3805–3812. [PubMed: 16597734]
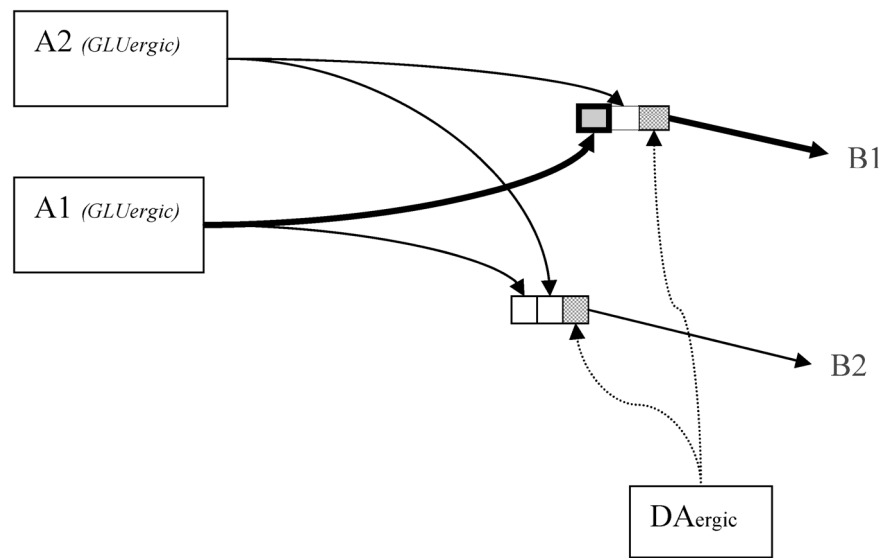
FIGURE 1.
Output neurons (B1, B2) in the dorsal and ventral striatum receive GLU inputs (A1, A2) originating in cortical and limbic regions. The informational nature of these GLU inputs varies according to the striatal target region.[5] For the sake of the illustration, we imagine that the inputs (**A**) carry information about the sensory environment, and the outputs (**B**) represent behavioral response tendencies. DA activity influences information flow through,[5,76] and plasticity (LTP and LTD)[8] of the currently active A→B GLU synapse (*bold square*). DA binding to receptors on the output cells (*stippled squares*) does not promote strengthening of nonactive GLU synapses (*other squares*). The diagram is simplified regarding the nature of the input–output connectivity. There is a large (up to 10,000 to 1) convergence of information from the cortex to striatum; that is, a given striatal neuron receives a large number of cortical and/or limbic GLU inputs (for an examination of corticostriatal mapping, see review; Ref. [77]). In addition, while dorsolateral striatal cells receive GLU input primarily from sensory-motor cortical regions, many striatal regions receive GLU inputs that carry information regarding sensory inputs, anticipated movements, expected outcomes, as well as information regarding appetitive and aversive valences of current inputs. Often, a single striatal neuron receives a convergence of such information, and responds to a conjunction of sensory, motor, and outcome–expectation conditions (see Ref. [5]; review).
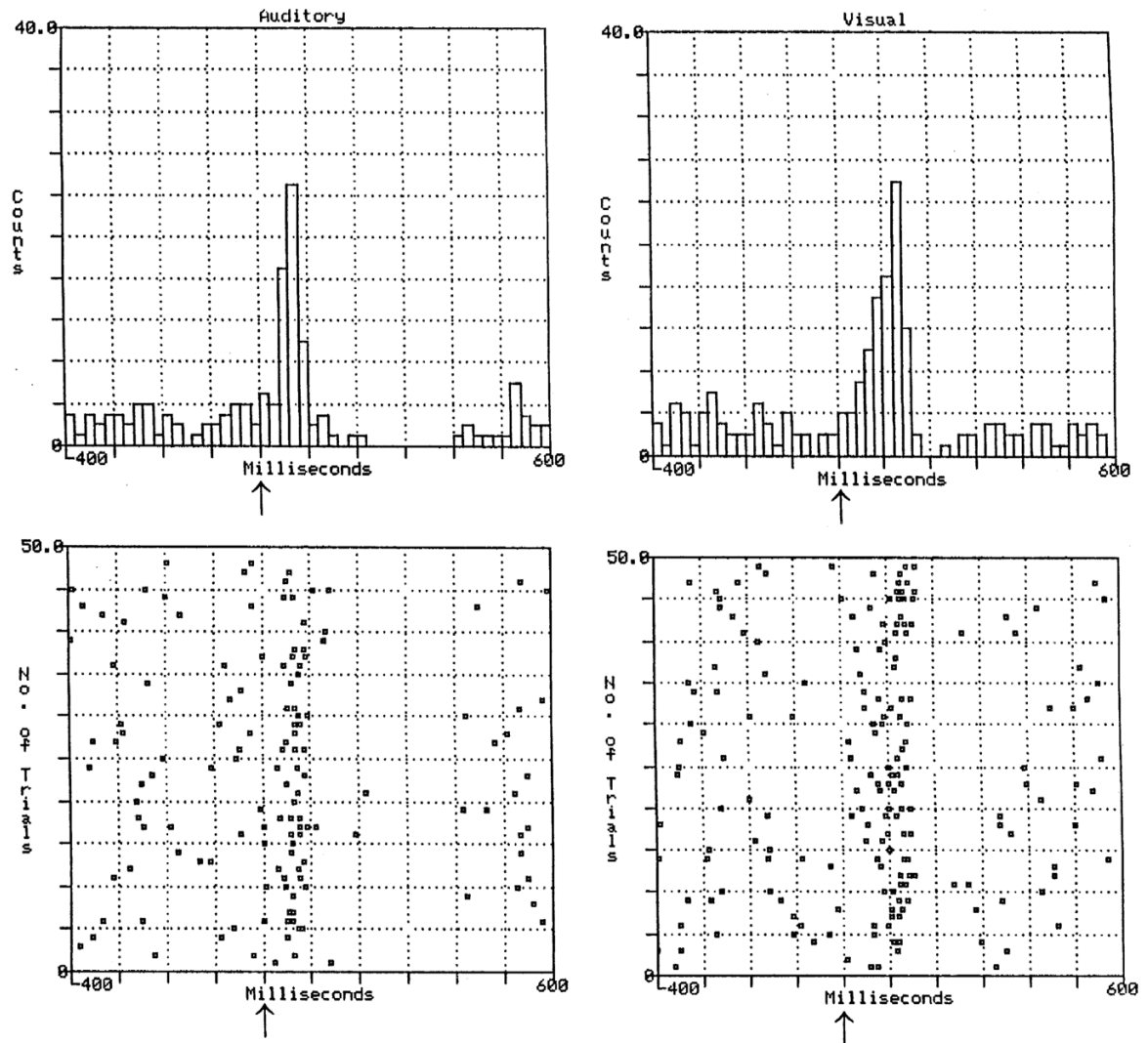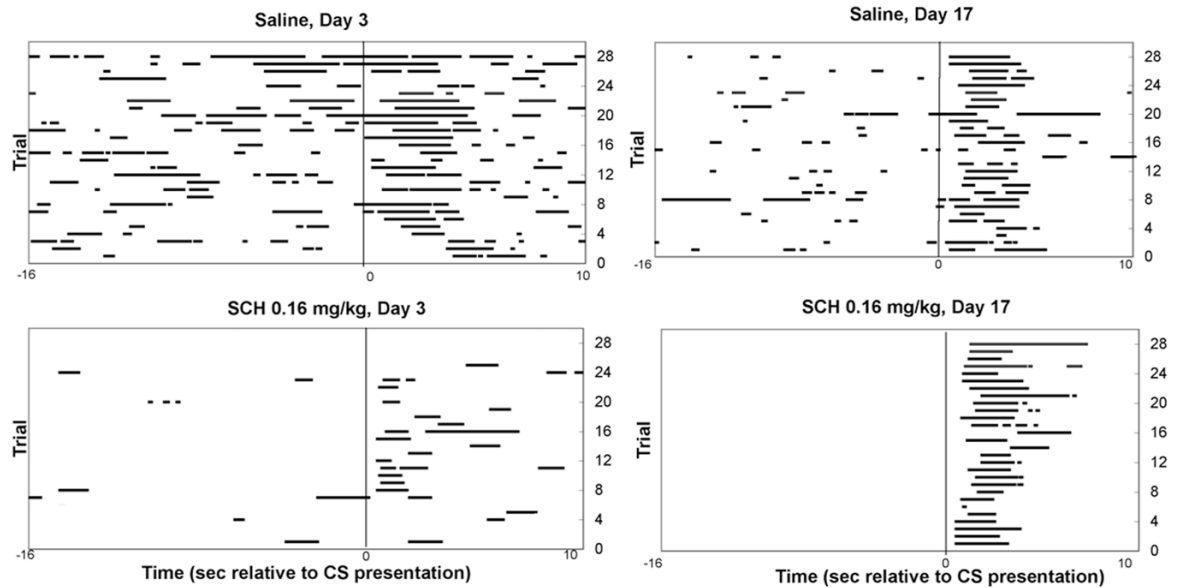
FIGURE 2.
The responses of an individual VTA dopamine neuron to an auditory click (*left panel*) and a light flash (*right panel*). X axis ranges from 400 msec before to 600 msec after the auditory or visual stimulus (*arrow*) of 1 msec duration (time 0).[24]

FIGURE 3.
(**A**) Left panels: Raster plot of head entries (horizontal lines) from −16 sec before to 10 sec after CS presentation (x axis). Successive trials 1–28 are represented from the bottom to the top of the y axis. On day 3 of training (after approximately 60 trials) systemic 0.16 mg/kg SCH23390 strongly suppresses head entries during the intertrial interval and in response to the CS (time 0). (**B**) Right panels: Separate groups of rats receive VEH or SCH23390 on day 17 of training. In animals receiving extended training prior to D1 antagonist challenge, the drug continues to suppress spontaneous head entries emitted during the ITI (−16 to 0), but does not affect the latency to respond to the CS.[17]
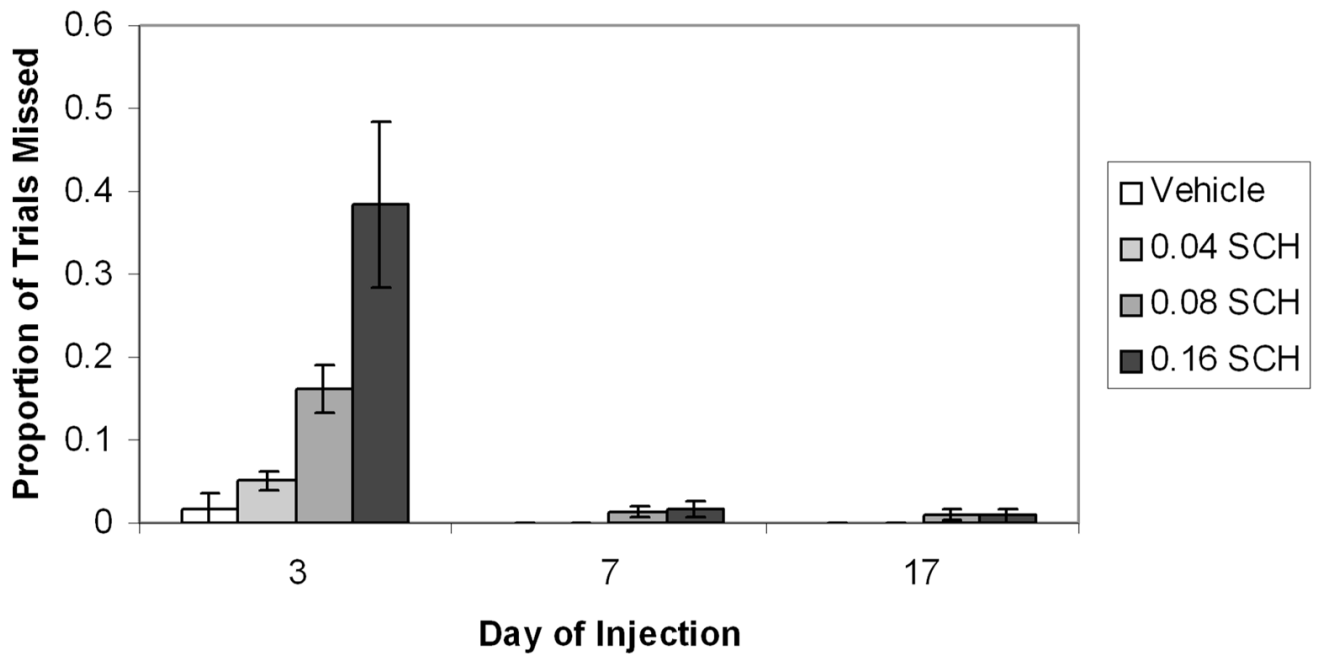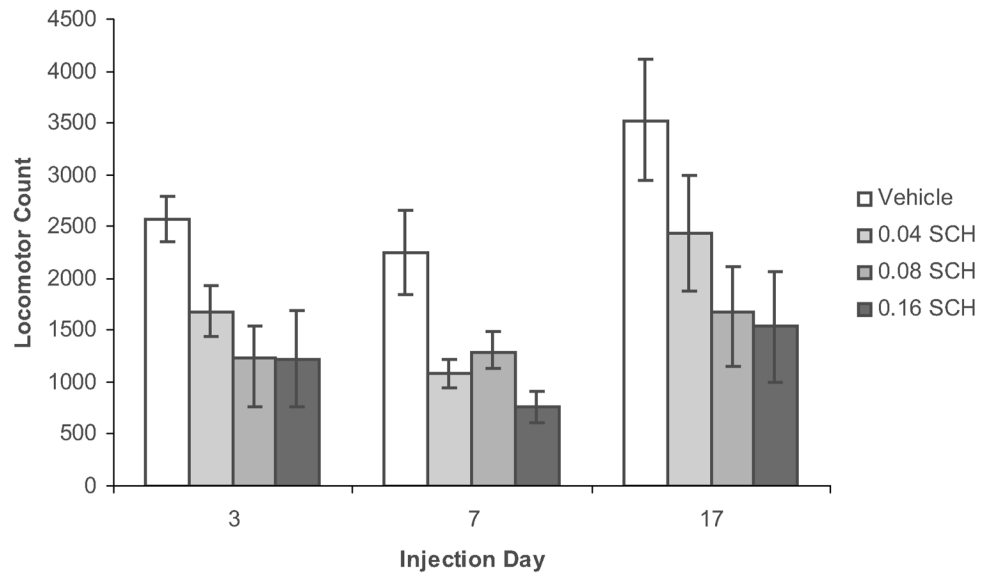
FIGURE 4.
Proportion of missed trials (i.e., trials for which latency to respond to the CS was > 10 sec) as a function of SCH23390 dose on test day 3, 7, and 17. *Bars* represent the standard error of the mean. On day 3, SCH23390 produced a dose-dependent increase in the proportion of missed trials ($P < 0.0005$). In contrast, animals that received SCH23390 on either day 7 or 17 of training showed no increase in missed trials. Note that animals in each group received only a single injection of SCH23390.[17]

FIGURE 5.
Two-way (Drug X Day) analysis of variance (ANOVA) conducted on days 3, 7, and 17 locomotor scores revealed a main effect of drug ($P < 0.000001$), a main effect of day ($P < 0.05$), and no interaction, showing that the drug produced similar locomotor suppression across different stages of learning.[17]
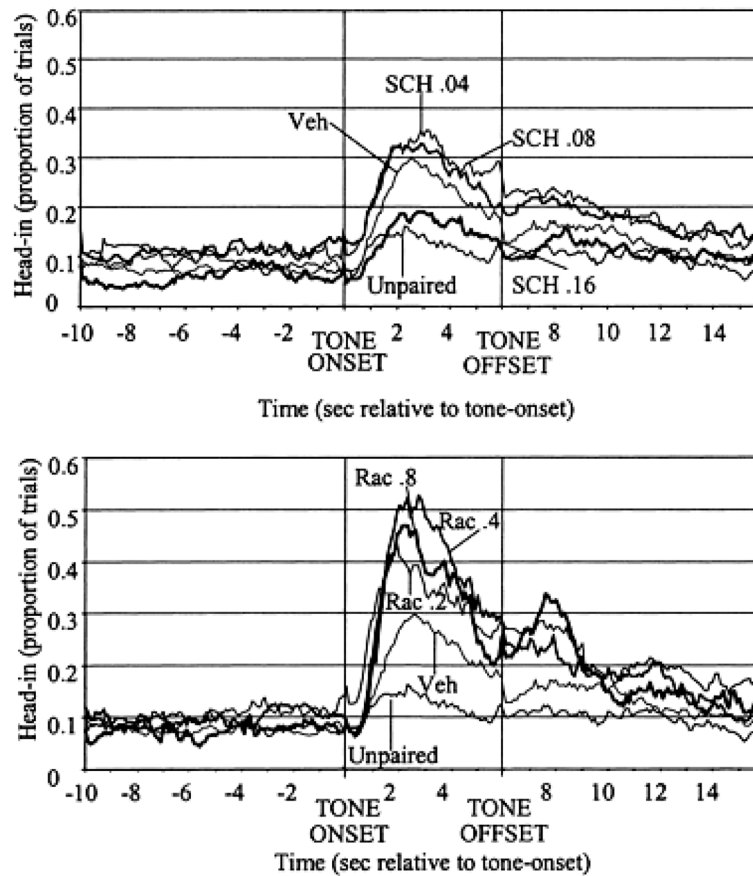
FIGURE 6.
Performance of each treatment group during day 19, the drug-free test session. The y axis shows mean proportion of trials during which rats' heads were in the food compartment for each successive 100 msec bin during the 10 sec before tone onset (−10 to 0), the presentation of the tone (0–6), and the 10 sec after tone offset (6–16). As can be seen, rats that had received UNPAIRED presentations of CS and food showed reduced CS period head-in durations compared to paired CS–food controls ($P < 0.01$). Animals that were under the influence of the highest SCH23390 dose during tone–food pairings also showed a reduced CS head-in probability ($P < 0.05$). Animals that were under the influence of RAC during the tone–food pairings showed increased CS head-in probabilities compared to vehicle controls on test day ($P < 0.01$).[10]