To appear in *The Handbook of Brain Theory and Neural Networks*, Second edition, (M.A. Arbib, Ed.), Cambridge, MA: The MIT Press, 2002.http://mitpress.mit.edu

The Roles of Dopamine

Jean-Marc Fellous and Roland E. Suri

The Salk Institute for Biological Studies
Computational Neurobiology Laboratory
10010 N. Torrey Pines rd,
La Jolla, CA 92037
Tel: 858-453-4100, x1618
Fax: 858-587-0417
fellous@salk.edu
suri@salk.edu

INTRODUCTION

Dopamine (DA) is a neuromodulator (see: NEUROMODULATION IN INVERTEBRATE NERVOUS SYSTEMS and SYNAPTIC CURRENTS, NEUROMODULATION AND KINETIC MODELS) that originates from small groups of neurons in the mesencephalon (the ventral tegmental area (A10), the substantia nigra (A9) and A8) and in the diencephalon (area A13, A14 and A15). Dopaminergic projections are in general very diffuse and reach large portions of the brain. The time scales of dopamine actions are diverse from few hundreds of milliseconds to several hours. We will focus here on the mesencephalic dopamine centers because they are the most studied, and because they are thought to be involved in diseases such as Tourette's syndrome, schizophrenia, Parkinson's disease, Huntington's disease, drug addiction or depression (see DISEASE: NEURAL NETWORK MODELS and (Tzschentke, 2001)). These centers are also involved in normal brain functions such as working memory, reinforcement learning, and attention. This article briefly summarizes the main roles of dopamine in particular with respect to recent modeling approaches.

BIOPHYSICAL EFFECTS OF DOPAMINE

The effects of dopamine on membrane currents and synaptic transmission are complex and depend of the nature and distribution of the postsynaptic receptors. At the single cell level, in the in vitro rat preparation, DA has been found to either increase or decrease the excitability of neurons, through the modulation of specific sets of sodium, potassium and calcium currents (see (Gulledge and Jaffe, 1998), and (Nicola et al., 2000) for reviews). While the exact nature of the modulation is still debated, it is likely to depend on the opposing contributions of the D1/D5 and D2/D3 family of dopamine receptors that are respectively positively and negatively coupled with adenylate cyclase. Studies in monkey cortical tissue showed that the D1/D5 family of receptor was 20-fold more abundant than the D2/D3 family, and that they were present distally in both pyramidal and non-pyramidal cells (Goldman-Rakic et al., 2000).

Dopamine modulates excitatory and inhibitory synaptic transmission. While the nature of neuromodulation of inhibitory transmission is still debated, it appears that in both the cortex and the striatum, D1 receptor activation selectively enhances NMDA but not AMPA synaptic transmission. Because of their voltage dependence, NMDA currents are smaller at rest than in a depolarized state when the postsynaptic cell is firing. Experimental and theoretical evidence suggest that the dopamine enhancement of NMDA currents may be used to

induce working memory-like (see below) bistable states in large networks of pyramidal neurons (Lisman et al., 1998).

In rats in vivo, stimulation of the ventral tegmental area or local application of dopamine decreases the spontaneous firing of the prefrontal cortex (Thierry et al., 1994), striatum and nucleus accumbens (Nicola et al., 2000), suggesting that dopamine may be able to control the levels of noise, and hence signal-to-noise ratios.

Given that dopamine modulation strongly depends on the particular distribution of D1/D5 and D2/D3 receptors and on the particular pattern of incoming synaptic transmission, the biophysical effects of dopamine on the intrinsic and synaptic properties is likely to differ from one neuron to the next, raising the intriguing possibility of the existence of several subclasses of neurons that differ only by their responses to this neuromodulator.

DOPAMINE LEVELS INFLUENCE WORKING MEMORY

Working memory refers to the ability to hold a few items in mind, with the explicit purpose of working with them to yield a behavior (see SHORT-TERM MEMORY). Typically, working memory tasks such as spatial delayed match-to-sample tasks consist in the brief presentation of a cue-stimulus (bright dot flashing once) in one of the 4 quadrants of a screen, followed by a delay period of several seconds, and by a test where the subject has to respond only if the test stimulus appears the same quadrant as the cue-stimulus. Single cells studies in monkeys revealed that some prefrontal cortical cells increased their firing rate during the delay period, when the stimulus is no longer present but when the animal has to remember its location in order to later perform the correct action. Both pyramidal cells and interneurons may present this property. The activity of these cells is stimulus dependent, so that only the cells that encode for the spatial location where the cue-stimulus occurred remain active during the delay period.

Local iontophoretic administrations of DA in the prefrontal cortex of monkeys performing a working memory task increase the cells' firing rate during the delay period, without increasing background noise, essentially increasing the signal-to-noise ratio during the task. There is however an optimal level of dopamine concentration above and below which working memory becomes impaired. Current theories propose that this effect is due to the enhancement by dopamine of excitatory inputs on pyramidal cells and interneurons observed in vitro. Because DA is more effective in facilitating excitatory transmission on pyramidal cells than on interneurons, intermediate levels of DA improves performance, while higher levels of DA recruits feed forward inhibition and decrease pyramidal cell outputs, therefore resulting in impairments in the task. Low levels of DA would not be sufficient in inducing excitatory facilitation, yielding a poor pyramidal cell output, and hence an impairment (Fig 1 and (Goldman-Rakic et al., 2000)). There has been a few attempts at modeling the neural substrate of working memory, but very little has yet been done to account for the role of dopamine (Tanaka, 2001).

DOPAMINE RESPONSES RESEMBLE REWARD PREDICTION SIGNAL OF TD MODEL

A large body of experimental evidence led to the hypothesis that Pavlovian learning depends on the degree of the unpredictability of the reinforcer (Dickinson, 1980). According to this hypothesis, reinforcers become progressively less efficient for behavioral adaptation as their predictability grows during the course of learning. The difference between the actual occurrence and the prediction of the reinforcer is usually referred to as the "error" in the reinforcer prediction. This concept has been used in the temporal-difference model (TD model) of Pavlovian learning (see REINFORCEMENT LEARNING IN MOTOR CONTROL). If the reinforcer is a reward, the TD model uses a reward prediction error signal to learn a reward prediction signal. The error signal progressively decreases and shifts to the time of earlier stimuli that predict the reinforcer. The characteristics of the reward prediction signal are comparable to those of anticipatory responses such as salivation in Pavlov's experiment.

The reward prediction error signal of the TD model remained a purely hypothetical signal until researchers discovered that the activity of midbrain dopamine neurons is strikingly similar to the reward prediction error of the TD model (Fig. 2A) (Montague et al., 1996; Schultz, 1998). Advances in reinforcement learning theories and evidence for the involvement of dopamine in sensorimotor learning and in cognitive functions lead to the development of the Extended TD model. The reward prediction error signal of the TD model by (Suri and Schultz, 1999) reproduces dopamine neuron activity in several situations: (1) upon presentation of unpredicted rewards, (2) before, during, and after learning that a stimulus precedes a reward, (3) when two stimuli precede a reward with fixed time intervals, (4) when the interval between the two stimuli are varied, (5) in the case of unexpectedly omitted reward, (6) delayed reward, (7) reward earlier than expected, (8) in the case of unexpectedly omitted reward-predictive stimulus, (9) in the case of a novel, physically salient stimulus that has never been associated with reward (see allocation of attention, below), (10) and for the blocking paradigm. To reach this close correspondence, three constants of the TD model were tuned to characteristics of dopamine neuron activity (learning rate, decay of eligibility trace, and temporal discount factor), some weights were initialized with positive values to achieve (9), and some *ad hoc* changes of the TD algorithm were introduced to reproduce (7) (see below).

In Pavlov's experiment, the salivation response of the dog does not influence the food delivery. The TD model is a model of Pavlovian learning and therefore computes predictive signals, corresponding to the salivation response, but does not select optimal actions. In contrast, instrumental learning paradigms, such as learning to press a lever for food delivery, demonstrate that animals are able to learn to perform actions that optimize reward. To model sensorimotor learning in such paradigms, a model component called the Actor is taught by the reward prediction error signal of the TD model. In such architectures, the TD model is also called the Critic. This approach is consistent with animal learning theory and was successfully applied to machine learning studies (see REINFORCEMENT LEARNING IN MOTOR CONTROL). Midbrain dopamine neurons project to the striatum and cortex and are characterized by rather uniform responses throughout the whole neuron population. Computational modeling studies with Actor-Critic models show that such a dopamine-like reward prediction error can serve as a powerful teaching signal for learning with delayed reward and for learning of motor sequences (Suri and Schultz, 1999). These models are also consistent with the role of dopamine in drug addiction and electrical self-stimulation (see below). Comparison of the Actor-Critic architecture to biological structures suggests that the Critic may correspond to pathways from limbic cortex via limbic striatum (or striosomes) to dopamine neurons, whereas the Actor may correspond to pathways from neocortex via sensorimotor striatum (or matrisomes) to basal ganglia output nuclei (see BASAL GANGLIA) (Fig. 2B). Whereas this standard Actor-Critic model mimics learning of sensorimotor associations or habits, it does not imply that dopamine is involved in anhedonia.

ALLOCATION OF ATTENTION

Several lines of evidence suggest that dopamine is also involved in attention processes. Although the firing rates of dopamine neurons can be increased or decreased for aversive stimuli, dopamine concentration in striatal and cortical target areas are often increased (Schultz, 1998). Both findings are not necessarily inconsistent since small differences in firing rates of dopamine neurons are hard to detect with single neuron recordings, and measurement methods for dopamine concentration have usually less temporal resolution than those of spiking activity of dopamine neurons. Furthermore, dopamine concentration is not only influenced by dopamine neuron activity but also by local regulatory processes. Slow changes in cortical or striatal dopamine concentration may signal information completely unrelated to reward. Otherwise, relief following aversive situations may influence dopamine neuron activity as if it were a reward, which would be consistent with opponent processing theories (See CONDITIONING). Allocation of

attentional resources seems to determine dopamine neuron activity in the situation when a reward is delivered earlier than usual. In contrast to any linear model, including the standard TD model, dopamine neuron activity is on base line levels at the time of the expected reward in this situation. This suggests that delivery of the reward earlier than usual seems to reallocate attentional resources through competitive mechanisms (Suri and Schultz, 1999).

Dopamine neurons respond to novel, physically salient stimuli even if the stimulus has never been associated to a reward (Schultz, 1998). In contrast to reward-predictive responses, for stimuli of equal physical salience, the increase due to novelty responses seems to be smaller and is followed by a pronounced decrease of neural activity below base line levels. (Brief and less pronounced decreases of dopamine neuron activity sometimes also occur after a response to a reward.) In contrast to responses to conditioned stimuli, novelty responses extinguish for repeated stimulus presentations. The characteristics of this novelty response is consistent with the TD model if certain associative weights are initialized with positive values instead of using initial values of zero (Suri and Schultz, 1999). Such initialization of initial weights with positive values was proposed in machine learning studies to stimulate exploration of novel actions. Simulation studies demonstrated that such a novelty bonus hardly influences slow movements of more than 100 msec duration because the effects of the two phases in the firing of dopamine neurons cancel out and the movement starts after the biphasic response. However, dopamine novelty responses may stimulate exploration for very brief actions, which may include saccades or allocation of attentional resources (Suri and Schultz, 1999).

Redgrave and collaborators (Redgrave et al., 1999) argued that the latency of dopamine responses is too short to be consistent with the hypothesis that dopamine is a reward prediction signal. Onsets of dopamine novelty responses as well as reward responses seem to occur just before the start of the saccade or during the saccade. The dopamine response will likely occur after the superior colliculus has detected a visual target but prior to the triggering (by collicular neurons) of the saccadic movement required to bring the target to the fovea. If it is assumed that the animal must execute a saccade to a visually presented stimulus before it can adequately assess its predictive value, the latency of dopamine response would be too short to signal reward. We argue against this view of Redgrave and colleagues. Neural activities in cortical and subcortical areas reflect the anticipated future visual image before a saccade is elicited (Ross et al., 2001). Therefore, these representations of future visual images may influence dopamine neuron activity as if the saccade had already been executed, and thus the dopamine response may start slightly before the saccade. The Extended TD model computes such predictive signals and uses them to select goal-directed actions in a cognitive task (Suri et al., 2001). According to this complex Actor-Critic model, the interactions between dopamine neuron activities (computed by Critic) and activities that reflect the preparation for intended actions (in Actor) select the actions that maximize reward predictions. The model evaluates the expected values of future actions, without necessarily executing them, in order to select the action with the optimal predicted outcome. The model selects the optimal action from such ‘action ideas’ or ‘imagined actions’. This optimal action is selected by assuming that dopamine neuron activity increases the signal-to-noise-ratio in target neurons. According to this advanced Actor-Critic model, dopamine improves focusing of attention to intended actions and selects actions. Since some neural activities anticipate the retinal images that result of saccades before these saccades are executed (Ross et al., 2001), animals may indeed use such predictive mechanisms for the selection of intentional saccades. Furthermore, similar internal mechanisms may bias intentional switching capabilities of the basal ganglia to facilitate the allocation of behavioral and cognitive processing capacity towards unexpected events (see BASAL GANGLIA and (Redgrave et al., 1999)). If we assume similar functions of dopamine for short-term memory, this model suggests that dopamine may select the items that should be kept in short-term memory and may also help to sustain their representation over time.

CONCLUSIONS

In vitro studies of the biophysical effects of dopamine demonstrate a wide range of dopamine effects on the intrinsic and synaptic properties of individual cells. *In vivo* studies suggest however that the main overall effect of dopamine may be to control noise levels and to selectively enhance the signal-to-noise-ratio of neural processing. This action may behaviorally lead to an improvement of working memory and to better selection of goal-directed actions. The TD model reproduces dopamine neuron activity in many behavioral situations and suggests that dopamine neuron activity code for an error in reward prediction. A complex TD model was described that solves cognitive tasks including goal-directed actions (also called planning or intentional) and attempts to reproduce the function of dopamine in attention and preparation processes.

FIGURE CAPTIONS

Fig. 1. Biphasic effects of Dopamine during a working memory task. The task consisted in the brief presentation of a cue (C), a delay of 3 seconds (D) and a response (R). Moderate levels of local application of SCH39166 (25 nA), a D1 receptor agonist, dramatically enhanced the activity of this cell, without significantly increasing its background activity (before cue). Higher levels of SCH39166 (75 nA) decreased the activity of this cell throughout the task. Histogram units are spikes/s. Figure adapted from (Goldman-Rakic et al., 2000).

Fig. 2. **A:** Prediction error signal of the TD model (left) similar to dopamine neuron activity (right) (figure adapted from (Suri and Schultz, 1998)). If a neutral stimulus A is paired with reward, prediction error signal and dopamine activity respond to the reward (before learning). After repeated pairings, the prediction error signal and dopamine activity are already increased by stimulus A and on baseline levels at the time of the reward (after learning). If the stimulus A is conditioned to a reward but is occasionally presented without reward, the prediction error signal and dopamine activity are decreased below baseline levels at the predicted time of reward (omitted reward). **B:** Interactions between cortex, basal ganglia, and midbrain dopamine neurons mimicked by Actor-Critic models. The limbic areas are proposed to correspond to the Critic and the sensorimotor areas to the Actor. The striatum is divided into matrisomes (semsorimotor) and striosomes (limbic). Limbic cortical areas project to striosomes, whereas neocortical areas chiefly project to matrisomes. Midbrain dopamine neurons are contacted by medium spiny neurons in striosomes and project to both striatal compartments. They are proposed to influence sensorimotor learning in the matrisomes (instrumental learning) and learning of reward predictions in the striosomes (Pavlovian learning). Striatal matrisomes inhibit the basal ganglia output nuclei Gpi/SNr and can elicit actions due to their projections via thalamic nuclei to motor cortical areas. Several additional functions of this architecture were proposed in (Suri et al., 2001).

REFERENCES

Dickinson A. (1980) Contemporary animal learning theory. Cambridge, UK: Cambridge University press.

Goldman-Rakic P.S., Muly E.C., 3rd, Williams G.V., 2000, D(1) receptors in prefrontal cells and circuits, Brain Res Rev, 31:295-301.

Gulledge A.T., Jaffe D.B., 1998, Dopamine decreases the excitability of layer V pyramidal cells in the rat prefrontal cortex, J Neurosci, 18:9139-9151.

Lisman J.E., Fellous J.-M., Wang X.-J., 1998, A role for NMDA-receptor channels in working memory, Nature Neuroscience, 1:273-275.

Montague P.R., Dayan P., Sejnowski T.J., 1996, A framework for mesencephalic dopamine systems based on predictive Hebbian learning, J Neurosci, 16:1936-1947.

Nicola S.M., Surmeier J., Malenka R.C., 2000, Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens, Annu Rev Neurosci, 23:185-215.

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

Ross J., Morrone M.C., Goldberg M.E., Burr D.C., 2001, Changes in visual perception at the time of saccades, Trends Neurosci, 24:113-121.

Schultz W., 1998, Predictive reward signal of dopamine neurons, J Neurophysiol, 80:1-27.

Suri R.E., Schultz W., 1998, Learning of sequential movements by neural network model with dopamine-like reinforcement signal, Exp Brain Res, 121:350-354.

Suri R.E., Schultz W., 1999, A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task, Neuroscience, 91:871-890.

Suri R.E., Bargas J., Arbib M.A., 2001, Modeling functions of striatal dopamine modulation in learning and planning, Neuroscience, 103:65-85.

Tanaka S., 2001, Computational approaches to the architecture and operations of the prefrontal cortical circuit for working memory, Prog Neuropsychopharmacol Biol Psychiatry, 25:259-281.

Thierry A.M., Jay T.M., Pirot S., Mantz J., Godbout R., Glowinski J., 1994, Influence of afferent systems on the activity of the rat prefrontal cortex: Electrophysiological and pharmacological characterization. In: Motor and Cognitive Functions of the Prefrontal Cortex (Thierry A.M., Glowinski J., Goldman-Rakic P.S., Christen Y., eds), pp 35-50. New York: Springer-Verlag.

Tzschentke T.M., 2001, Pharmacology and behavioral pharmacology of the mesocortical dopamine system, Prog Neurobiol, 63:241-320.

Fig 1

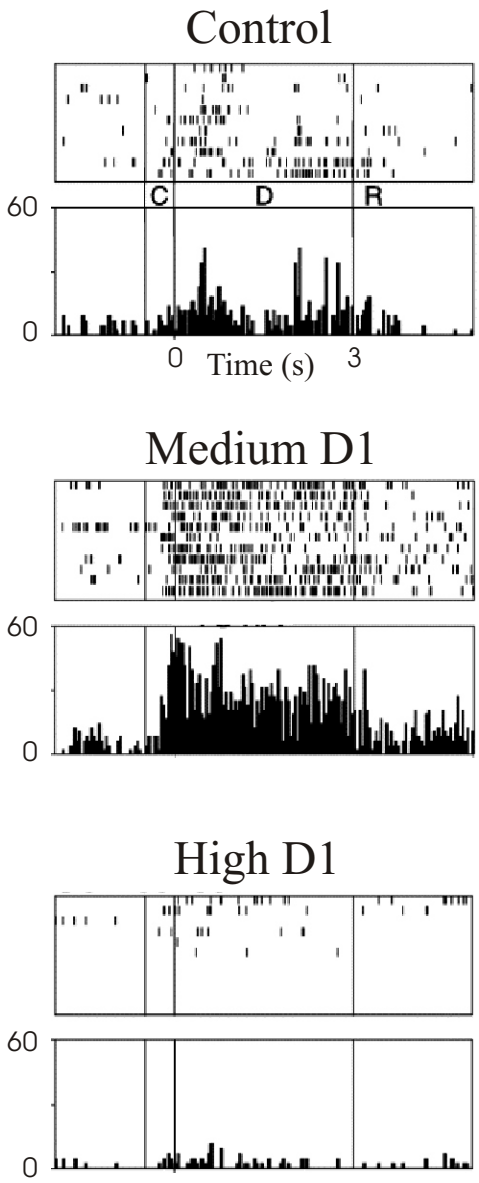
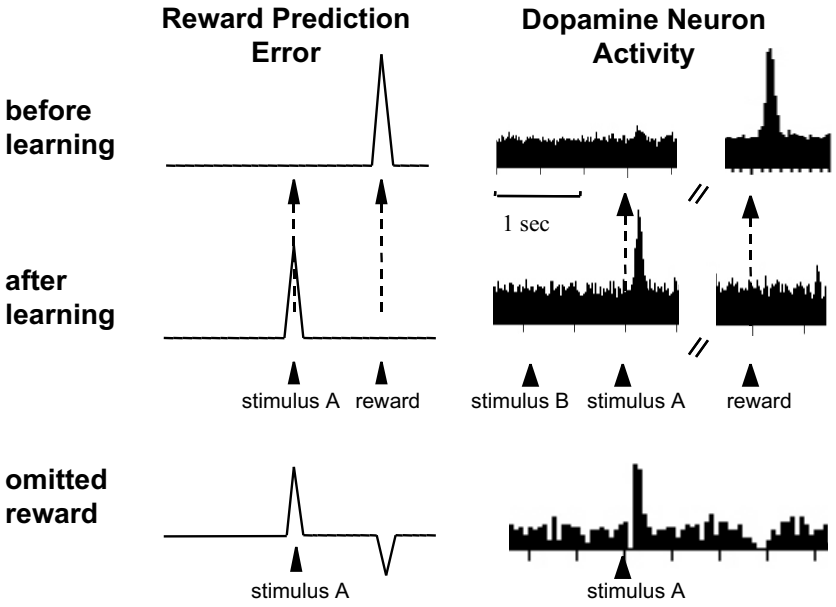
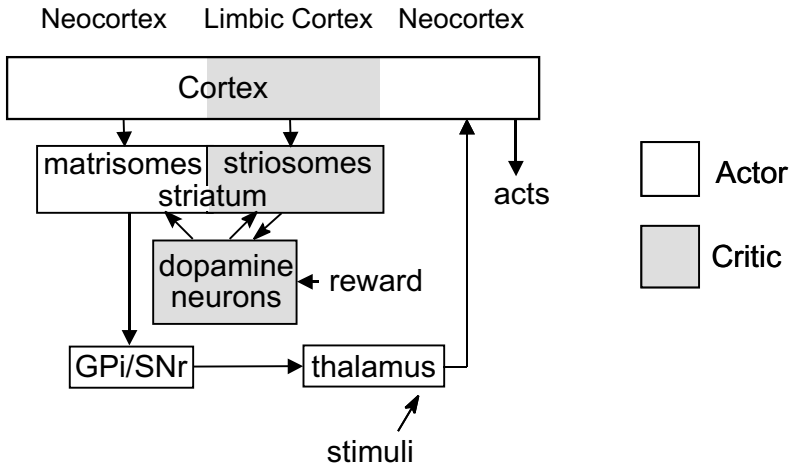


Fig 2

A


Reward Prediction Error
Dopamine Neuron Activity
before learning
after learning
omitted reward
1 sec
stimulus A
reward
stimulus B
stimulus A
reward
stimulus A
stimulus A

B


Neocortex
Limbic Cortex
Neocortex
Cortex
matrisomes
striosomes
striatum
acts
dopamine neurons
reward
GPi/SNr
thalamus
stimuli
Actor
Critic