# A computational substrate for incentive salience

## Samuel M. McClure[1], Nathaniel D. Daw[2] and P. Read Montague[1]

[1]Center for Theoretical Neuroscience, Human Neuroimaging Laboratory, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030, USA
[2]Computer Science Department, Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213, USA

**Theories of dopamine function are at a crossroads. Computational models derived from single-unit recordings capture changes in dopaminergic neuron firing rate as a prediction error signal. These models employ the prediction error signal in two roles: learning to predict future rewarding events and biasing action choice. Conversely, pharmacological inhibition or lesion of dopaminergic neuron function diminishes the ability of an animal to motivate behaviors directed at acquiring rewards. These lesion experiments have raised the possibility that dopamine release encodes a measure of the incentive value of a contemplated behavioral act. The most complete psychological idea that captures this notion frames the dopamine signal as carrying 'incentive salience'. On the surface, these two competing accounts of dopamine function seem incommensurate. To the contrary, we demonstrate that both of these functions can be captured in a single computational model of the involvement of dopamine in reward prediction for the purpose of reward seeking.**

Many aspects of behavior are dedicated to identifying, judging and pursuing objects and goals. It has long been recognized that the mesencephalic dopamine system is centrally involved in these processes for a range of behaviors, over and above its role in purely motor functions. One classic attempt to account for dopamine-mediated events in reward and motivation was the anhedonia hypothesis of dopamine dysfunction [1], which proposed that dopamine release was the hedonic signal of the brain that represented the pleasure associated with a primary reward. According to this hypothesis, deficits caused by interference with dopamine followed from the lack of this hedonic signal. A variety of subsequent experimental results have challenged this initial formulation. In particular, manipulations leading to reduced dopamine-mediated function do not affect hedonic responses to primary rewards [2], and increased activity of dopaminergic neurons is often tied to events that precede reward consumption [3]. In light of these data, several theories have been posited to replace the anhedonia hypothesis, all of which implicate dopamine in reward anticipation or reward seeking [2,4−6]. These theories are generally depicted as being inconsistent with one another.

On the contrary, we find that the degree of overlap of the experimental findings represents a tremendous improvement in our integrated understanding of dopaminergic neuron function.

We begin with a description of an experiment that captures the core of the incentive salience hypothesis of dopamine function. This description is then seated within the prediction error theory of dopamine function, which suggests that phasic changes in dopaminergic cell firing encode an error in the prediction of future reward (Eqn 3) [3,4]. This approach offers an extendable, mathematical model that captures the concept of incentive salience. We also show how this model explains a group of earlier behavioral experiments that reveal clearly the dual function of dopamine release as both a learning signal and an action bias.

## The incentive salience hypothesis

The incentive salience hypothesis is grounded in a growing history of behavioral effects that accrue in response to manipulation of dopaminergic neuron function. One singular finding is that dopamine receptor antagonism does not change the appetitive value of rewards, but appears to inhibit the ability to initiate actions necessary to acquire the rewards [2,5−7]. These findings have led to the idea that dopamine receptor antagonism selectively inhibits the capacity to initiate reward seeking actions, but has no effect on the value of the reward to the animal [2,5−7]. That is, the internal valuation of the reward by the animal is not changed, but its capacity to act on that valuation is inhibited. This view of dopamine function is supported by an enormous body of literature, which we do not seek to summarize. In our opinion, the most fruitful summary of these effects is captured in the idea of incentive salience as proposed by Berridge and Robinson [2] (but see Refs [5,6] for modifications of their basic proposal). The idea of incentive salience is that dopamine release assigns incentive value to objects or behavioral acts. These assigned values are then available to be used by some action selection system that makes more valuable actions more likely. Accordingly, antagonism of dopamine receptor function does not influence the assignment of value, but does inhibit the use of these values in choosing actions.

An experiment by Ikemoto and Panksepp [8] (Fig. 1) highlights the essential features of the incentive salience hypothesis. Rats were trained to traverse a one-arm maze
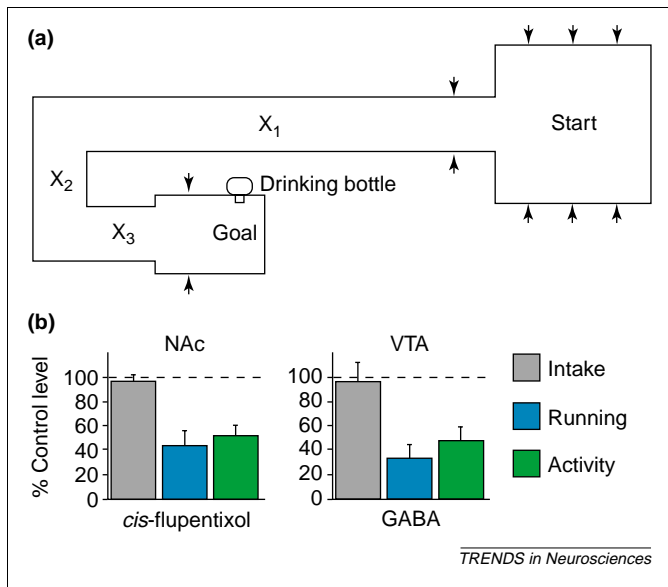
---

**Fig. 1**. The incentive salience hypothesis. (a) Rats were trained to traverse a one-armed maze to obtain sucrose solution at the end. Photosensors (arrows) were used to determine running speed [blue bars in (b)], in addition to the baseline level of movement in the start box [green bars in (b)], while access to the runway was blocked by a door. X₁, X₂ and X₃ represent intermediate states in the model, as shown in Fig. 3. (b) After training, dopaminergic neuron activity was reduced either by application of the dopamine receptor antagonist *cis*-flupentixol in the nucleus accumbens (NAc) or by injection of GABA into the ventral tegmental area (VTA). Both manipulations reduced the ability of the rats to initiate the running needed to acquire the sucrose solution ($P < 0.01$; loss of ability to ascribe incentive salience), while leaving the volume of sucrose they consumed unaffected (gray bars). Using data from Ref. [8].

to obtain access to sucrose solution at the end (Fig. 1a). After training, investigators administered either saline or one of three pharmacological agents into specific areas in the brain. Two of the manipulations degraded dopaminergic neuron function. Both of these, injection of GABA

into the ventral tegmental area and injection of the dopamine receptor agonist *cis*-flupentixol into the nucleus accumbens, caused dramatic behavioral deficiencies (Fig. 1b). The baseline levels of movement of the rats outside the context of the task were dramatically decreased. Furthermore, their running speeds within the maze were significantly less than those of saline controls. However, the treated rats maintained the same desire for the sucrose solution: when the investigators moved them to the bottle, they drank normal volumes of liquid during a 30 s reward period (Fig. 1b, gray bars). This latter observation demonstrates that dopaminergic neuron antagonism did not affect the motivation for the reward of the rats.

These results have been interpreted to mean that dopamine does not encode the pleasure associated with rewards, as the anhedonia hypothesis suggests, but instead enables reward-seeking behaviors. Changes in dopamine-mediated activity are appropriately timed for such a role: increased activation precedes reward-motivated actions, instead of being linked to the time of reward consumption [4]. Moreover, behavioral measures of affective reactions to primary rewards are unaltered by dopaminergic neuron antagonism [2]. Based on these behavioral data, Berridge and Robinson suggest that dopamine is responsible for assigning 'incentive salience' to objects and behavioral acts. Incentive salience maps 'liked' objects or acts to 'wanted' objects or acts (i.e. objects or acts that an animal will 'work to acquire') [2,6]. In their words, dopamine 'transforms the neural representation of a stimulus into an object of attraction that animals will work to acquire', and this attraction is transferred to conditioned stimuli associated with reward [2]. This mapping is illustrated qualitatively in Fig. 2a.
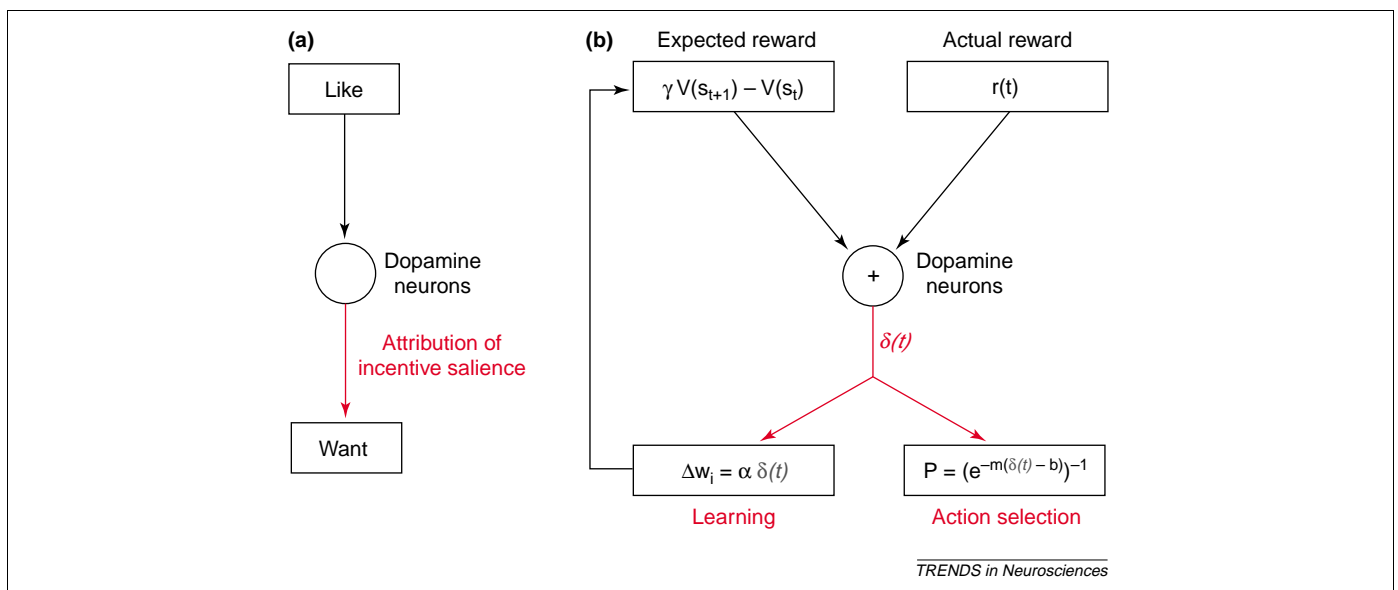


**Fig. 2**. Theories of dopamine function. (a) The incentive salience hypothesis. Dopamine is proposed to be responsible for converting an object or action that is 'liked' into one that is 'wanted'. In other words, dopamine ascribes incentive salience to objects and actions. This process is necessary to motivate actions aimed at acquiring rewards. (b) The prediction error hypothesis. Changes in dopaminergic neuron firing rate are hypothesized to encode a prediction error ($\delta$) signal as part of a computational system dedicated to seeking and predicting rewards (temporal-difference model). Inputs from the world (states, $s_t$) are used to generate an internal estimate of a value function ($V$) according to their learned weights ($w_i$). The temporal derivative of $V$ is then compared with the current reward received from the world ($r$) to generate the prediction error signal, $\delta$ (through weighting by a factor $m$ and subject to dopamine blockade through $b$). This error signal is used for two purposes: (1) as a learning signal to improve estimates of $w_i$, and (2) to bias action selection. It is this latter function, the conversion of learned incentive value into a probability of action, that is equivalent to the incentive salience hypothesis.

Several other theories of dopamine function are similar to the incentive salience hypothesis [5,6,9]. Although these theories differ in important ways, they all agree on the singular principle that dopamine function is causally located between the identification of a potential future reward and the generation of action to pursue it. Recent computational accounts of the dopamine system [4,10] (see also Refs [11,12]) wholly support this hypothesis; furthermore, the computational models extend the psychological descriptions by assigning parameters to key interactions in sets of simple equations. As detailed below, these models are commonly identified by their suggestion that changes in dopamine cell firing encode a prediction error in the amount of expected future reward.

### The prediction error theory of dopaminergic neuron firing and its relationship to a model of incentive salience

Computational accounts of dopamine function start by identifying the problems that neural reward systems must solve. One crucial challenge faced by the nervous system is how to choose actions to obtain food, sex and other natural rewards required for survival. This formidable problem is neglected by accounts such as the incentive salience hypothesis, which leave off after the identification of a goal. Various solutions to the problem of planning actions to obtain future reward have been proposed in the computer science community. A relatively simple approach is known as temporal difference (TD) learning [13,14]. It is the algorithm that underlies the prediction error explanation of dopaminergic neuron responses.

The TD algorithm aims to learn an estimate of a value function, $V^*$. The function relates the situation at a particular time, $s_t$, to the expected, time-discounted sum of rewards (idealized as numeric measures $r$ of received utility) that can be earned into the infinite future. For simplicity, we use an arbitrary discrete timescale ($t$, $t + 1$ and so on) with no specific relation to continuous real-time measurements (e.g. in minutes or seconds).

$$V^*(s_t) = E[r_t + \gamma\, r_{t+1} + \gamma^2 r_{t+2} + \gamma^3 r_{t+3} + \ldots] \qquad \text{[Eqn 1]}$$

The expectation is over-randomness in reward delivery and state transitions, and $0 < \gamma < 1$ is a discounting parameter. Were this function known, optimal decision-making would simply amount to choosing those actions that lead to the highest-valued states.

Although Eqn 1 is an infinite sum (so a single instance of it can never be fully observed), it can be rewritten in a recursive form [15] that is more suitable for learning:

$$V^*(s_t) = E[r_t + \gamma\, V^*(s_{t+1})] \qquad \text{[Eqn 2]}$$

The TD algorithm uses this relationship to refine successively an estimate of $V^*$, which we call $V$, using only finite chunks of experience. Equation 2 can be rearranged into a measure, $\delta$ of the extent to which the value estimates corresponding to a pair of successively observed states and an observed reward are mutually consistent:

$$\delta(t) = r_t + \gamma\, V(s_{t+1}) - V(s_t) \qquad \text{[Eqn 3]}$$

This can be used as an 'error signal' to nudge $V(s_t)$ towards a better estimate. For example, unexpected rewards or increases in $V(s_{t+1})$ will produce positive $\delta(t)$, an indication that $V(s_t)$ was too low; conversely, prediction of too much reward leads to negative prediction error. Besides improving an estimate of future reward, $\delta$ can be used to bias decision making towards actions that lead to better-than-expected reward (high $\delta$; see below).

A vast amount of neuronal recording data are explained by hypothesizing that the rate of dopamine neuron spiking encodes a TD prediction error signal, $\delta(t)$ [3,4,10]. Additionally, the model helps explain dopamine concentrations in the striatum during intracranial electrical stimulation [16,17]. The model also provides an accurate hypothesis for interpreting activity changes in human brain reward structures measured using functional magnetic resonance imaging [18–21].

In the TD model, dopamine serves two purposes (Fig. 2b). First, it is used as a learning signal for $V$, and is therefore hypothesized to be required for learning to predict future rewards. Second, dopamine release biases action selection towards situations predictive of reward. There are several schemes by which value estimates and the error signal can contribute to decision making, and the particular method the brain uses is unclear. Dopamine can affect action selection indirectly through its role in learning value predictions or, in some formulations (e.g. Ref. [22]), by directly modifying action selection weights: positive error increases the value attributed to some state or action, making it more likely to be chosen in the future. Dopamine release might also have more direct, immediate effects on action choice [4,11,12,23]: for example, as the prediction error reports whether pay-off is better or worse than expected, it can indicate whether to continue a course of action or to try something else. We use a simple action selection model [24] that incorporates both direct and indirect effects of dopamine on action choice. The specific details of the action selection function are not crucial for making our point. Instead, the important point is that, in both TD theories and the incentive salience hypothesis, increased dopamine activation has the role of increasing the likelihood of choosing some action that leads to reward.

To clarify, we propose that the concept of incentive salience is the expected future reward (Eqn 1). In addition, we propose that the role of dopamine in learning to attribute such expectations to situations that are predictive of reward (Eqn 3) and in biasing action selection towards such situations (e.g. Eqn 4) serve as the formal counterpart to the ideas of Berridge and Robinson [4] about the role of dopamine in attributing and using incentive salience.

### A TD learning account of incentive salience effects

To illustrate the integrated view outlined above, we adapted a TD model to the task studied by Ikemoto and Panksepp (Fig. 3). The world of the rat was modeled as consisting of five possible states: the start box, the goal box and three positions within the maze. The model stores an estimate $V(s)$ of the value of each state. From any position within the maze, the modeled rat could either move
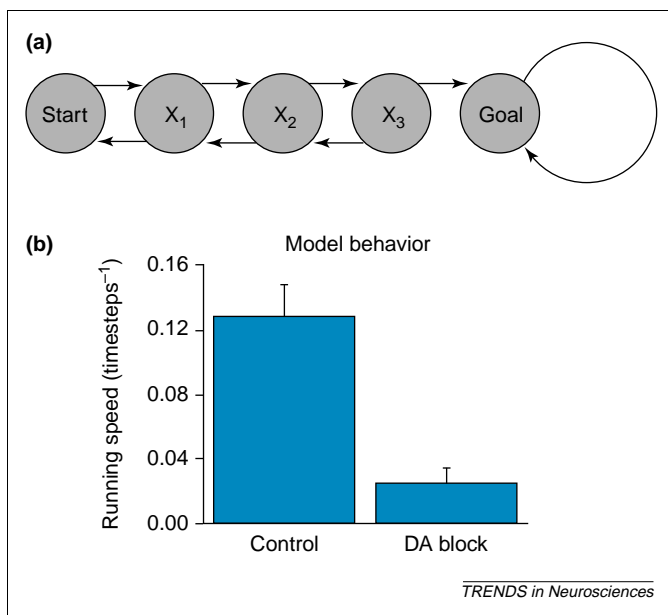
**Fig. 3**. Temporal difference model captures rat behavior on maze. (a) The maze from Fig. 1 is represented as consisting of five states. Virtual rats could either be in the start box, goal box or one of three intermediate states ($X_1$–$X_3$). Movements between states were determined by considering candidate transitions, which produce an error signal ($\delta$), and a probability of taking the considered action. (b) When virtual dopamine (DA)-receptor antagonists are added (Eqn 4 of main text; $b = 1$), the speed with which the rat traverses the maze is significantly decreased, as seen in real animals.

towards or away from the goal. Upon taking one of these actions, the rat obtains the reward of that state and is then allowed to continue until it reaches the goal box, after which it is returned to the start. The reward delivered was 1 for the goal box and 0 for all other states.

Actions were chosen based on the predicted values of their successor states, using the decision model of Egelman *et al.* [24]. The model sequentially evaluates a series of possible actions until it decides to accept one. A prediction error signal (Eqn 3) is computed based on the successor state of each contemplated action, and the action is accepted with probability $P$ given by a softmax function [14]:

$$P = (1 + e^{-m(\delta(t)-b)})^{-1} \qquad \text{[Eqn 4]}$$

where $m$ is a scaling constant and dopamine receptor inhibition was modeled by $b$, a constant subtracted from the dopamine signal, $\delta(t)$. If an action is not selected at any given moment according to $P$, then time is incremented forwards and another action is considered. Here, dopamine has direct effects on action selection because it reports the relative usefulness, $\delta(t)$, of contemplated actions. Its indirect effects come from its role in learning the values $V$, which underlie these reports. This learning is accomplished by updating the stored value estimates following each action, according to

$$V(s_i) \leftarrow V(s_i) + \alpha(\delta(t) - b) \qquad \text{[Eqn 5]}$$

where $\alpha$ is the learning rate and $s_i$ is the state that has just been left.

After the value estimates were learned, the effect of high levels of dopamine receptor inhibition was determined

(Fig. 3). We model dopamine receptor antagonism as a constant decrease in the modeled dopamine-mediated signal by subtracting a constant baseline $b$ from the error signal $\delta$ in Eqns 4 and 5. This discourages motivated behavior both through the indirect and direct effects of dopamine; notably, it directly reduces the probability of accepting any action (Eqn 4), thereby increasing the time required to reach the goal (reduced running speed). This is precisely the effect detailed by Ikemoto and Panksepp [8] (Fig. 1). In general, under dopamine-receptor inhibition, the direct dopamine-mediated report of the desirability ('incentive salience') of contemplated actions is suppressed, and the modeled rat is unable to engage the actions.

## A low concentration of dopamine-receptor antagonists causes gradual extinction

A subtler pattern of deficits occurs on the same task when lower concentrations of dopamine-receptor antagonists are administered. Specifically, behavioral changes do not occur immediately after drug delivery, but appear only through repeated exposure (but see Ref. [6]). This slow unlearning has been likened to the extinction that occurs when a conditioned stimulus is unpaired from reward delivery. Ikemoto and Panksepp [5] also cite this result as counter evidence to the incentive salience hypothesis, which has no provision (apart from some informal discussion of incremental attribution of incentive salience by a 'boosting' process) for delayed effects. However, when incentive salience is placed within a formal computational framework of reward learning and action selection, then the phenomenon is easily explainable as gradual extinction of the value estimates $V$.

Wise and colleagues [25] tested rats under low concentrations of the dopamine-receptor antagonist pimozide ($0.5\ \text{mg kg}^{-1}$ and $1.0\ \text{mg kg}^{-1}$), delivered systemically, on a one-armed maze (as in Ikemoto and Panksepp's [5] experiment; Fig. 1a). Treated animals were not initially worse at the task, but their running times slowed progressively over the course of a second day of testing, one week later (Fig. 4a). This behavior mirrored the effects of extinction in another group of rats from which food was withheld at the goal (Fig. 4a, NR condition). The fact that running speeds were not immediately affected argues against an explanation of the results based on impeded motor function.

Experience-dependent extinction of response is a direct consequence of the prediction error hypothesis (Fig. 4b), as a reduction in the TD error signal caused by dopamine-receptor inhibition will extinguish value predictions $V$ in much the same way as reward omission. If the animal has learned that the goal box is associated with food, no prediction error should occur when the box is reached and the food is delivered normally. Thus, dopamine neurons would be expected to spike at their baseline rate. However, owing to inhibition of postsynaptic dopamine receptors, this baseline response would produce a below-baseline rate of postsynaptic receptor activation, equivalent to negative prediction error signal at the time of the reward. This is the same neural response as is seen when rewards are withheld [10]. In terms of the model, the negative
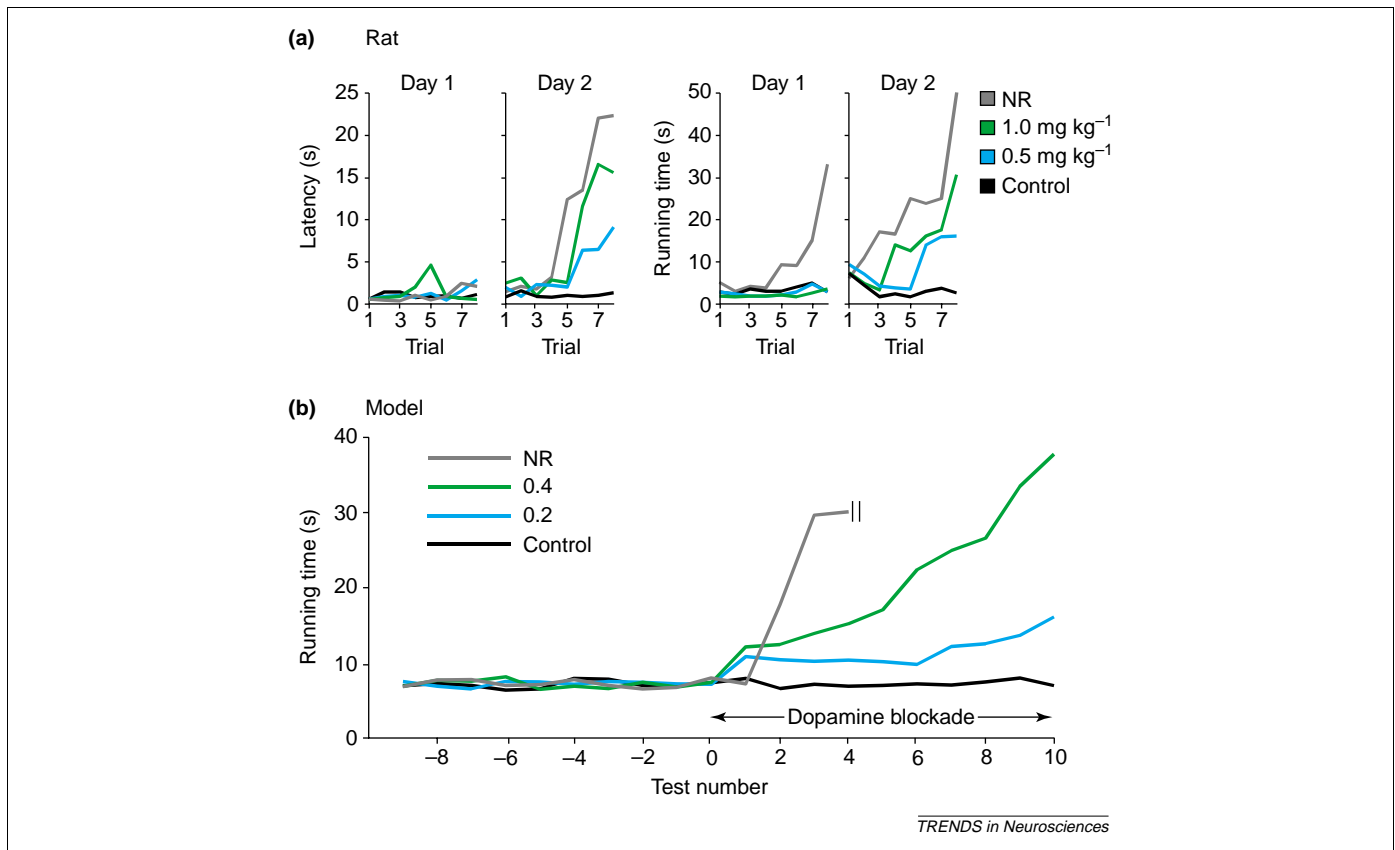
**Fig. 4**. Prediction error hypothesis accounts for extinction of maze running. (a) When rats are given low concentrations of the dopamine-receptor antagonist pimozide after learning to run a maze for a food-pellet reward, behavioral effects are not immediately seen. Instead, both latency and running speed remain unaffected during repeated trials on the first training day (Day 1). When tested again one-week later (Day 2), deficits become apparent and become stronger in an experience-dependent manner. Rats show a slow increase in latency before entering the maze, and decreased running speed through the maze [0.5 mg kg$^{-1}$ (blue) and 1.0 mg kg$^{-1}$ (green) pimozide conditions] in a manner that parallels the effect of extinction (NR condition, gray). Using data from Ref. [25]. (b) This effect is captured by the prediction error hypothesis as a slow decrease in estimated values ($V$) of each state in the maze. The different concentrations of pimozide were captured in the model as different scalar values ($b$) subtracted from the dopamine signal, $\delta$ [$b = 0.2$ (green) and $b = 0.4$ (blue), arbitrary units]. Extinction (NR, gray) is captured by setting to zero the reward value of arriving at the goal box. The model captures the behavioral data in a manner difficult to achieve using the incentive salience hypothesis alone.

prediction error signal should cause a decrease in the weights that underlie the estimated values of the goal box and the states that precede it (Eqn 5), causing an experience-dependent decrease in the action, as seen by Wise and colleagues.

Thus, according to the model, in situations when dopamine-receptor antagonism is weak enough only to disrupt minimally actions through the 'direct' action of dopamine (in the evaluation of the value of a contemplated action outlined in Eqn 4), 'indirect' effects of the antagonism will nonetheless build up gradually (through the learning of value weights outlined in Eqn 5). Low concentrations of pimozide should cause a progressive decrease of the incentive salience (i.e. expected future value) attributed to situations predictive of reward, and thus progressively disrupt actions that seek out such situations. Our formal computational account of the attribution and use of incentive salience in a broader action selection system significantly clarifies this point; according to the previous informal account, it was unclear whether the results were to be expected at all.

## Concluding remarks

We have proposed a mapping between a psychological theory of the role of dopamine in reward and motivation, and a more formal computational theory of how the neurotransmitter is involved in a larger system for choosing optimal actions under the motivation of prediction errors. Specifically, we have identified the concept of incentive salience [2] with the computational notion of expected future value, and have suggested that the TD theory of learning future values formalizes Berridge and Robinson's ideas about attributing incentive salience through a 'boosting' process. One major advantage of this maneuver is that the elements of the incentive salience idea are now parameterized and are quantitatively testable against detailed experimental data. In fact, there are already extant experimental data that immediately suggest extensions to the model as described.

In experiments by Ikemoto and Panksepp [8], animals with dopamine-receptor antagonism in the nucleus accumbens are poorer at traversing the maze but not at consuming the reward (e.g. at licking) when compared with controls. The basic TD framework does not distinguish these actions: both are motor actions that must be taken to obtain the primary reward. This point seems equally problematic for Berridge and Robinson's original incentive salience formulation [2], which assumes that reward consumption is a measure of incentive salience. The result makes more sense in the context of Ikemoto and Panksepp's [5] similar hypothesis, in which dopamine underlies appetitive behaviors such as approach, but not

consummatory behaviors such as licking [5]. TD models are easily extended to incorporate an assumption that consumptive actions have an automatic or habitual quality that makes them insensitive to dopamine-receptor inhibition [11,12,26,27]. In particular, Dayan and colleagues have provided such an extension by modifying a TD model to include such features of contemporary theories of instrumental conditioning as habitual versus goal-directed responding [11,12].

We have shown that various theories of dopamine function agree to a much greater extent than has been previously appreciated. In many ways, the appearance of discrepancies resulted from the different languages used by in the computationally and psychologically based arguments. However, when the different semantics are resolved, commonalities emerge that reveal real progress in our understanding of dopamine function.

### References
1 Wise, R.A. (1982) Neuroleptics and operant behavior: the anhedonia hypothesis. *Behav. Brain Sci.* 5, 39–87
2 Berridge, K.C. and Robinson, T.E. (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Rev.* 28, 309–369
3 Schultz, W. (1998) Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27
4 Montague, P.R. *et al.* (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947
5 Ikemoto, S. and Panksepp, J. (1999) The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res. Rev.* 31, 6–41
6 Salamone, J.D. and Correa, M. (2002) Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav. Brain Res.* 137, 3–25
7 Salamone, J.D. *et al.* (1997) Behavioral functions of nucleus accumbens dopamine: empirical and conceptual problems with the anhedonia hypothesis. *Neurosci. Biobehav. Rev.* 21, 341–359
8 Ikemoto, S. and Panksepp, J. (1996) Dissociations between appetitive and consummatory responses by pharmacological manipulations of reward-relevant brain regions. *Behav. Neurosci.* 110, 331–345
9 Redgrave, P. (1999) Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci.* 22, 146–151
10 Schultz, W. *et al.* (1997) A neural substrate of prediction and reward. *Science* 275, 1593–1599
11 Dayan, P. (2002) Motivated reinforcement learning. In *Advances in Neural Information Processing Systems* (14) (Dietterich, T. *et al.*, eds), pp. 11–18, MIT Press
12 Dayan, P. and Balleine, B.W. (2002) Reward, motivation, and reinforcement learning. *Neuron* 36, 285–298
13 Sutton, R.S. and Barto, A.G. (1998) *Reinforcement Learning*, MIT Press
14 Dayan, P. and Abbott, L.F. (2001) *Theoretical Neuroscience*, MIT Press
15 Bellman, R.E. (1957) *Dynamic Programming*, Princeton University Press
16 Garris, P.A. *et al.* (1999) Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. *Nature* 398, 67–69
17 Kilpatrick, M.R. *et al.* (2000) Extracellular dopamine dynamics in rat caudate-putamen during experimenter-delivered and intracranial self-stimulation. *Neuroscience* 96, 697–706
18 Berns, G.S. *et al.* (2001) Predictability modulates human brain response to reward. *J. Neurosci.* 21, 2793–2798
19 Pagnoni, G. *et al.* (2002) Activity in human ventral striatum locked to errors of reward prediction. *Nat. Neurosci.* 5, 97–98
20 McClure, S.M. *et al.* (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38, 339–346
21 O'Doherty, J. *et al.* (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337
22 Houk, J.C. *et al.* (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In *Models of Information Processing in the Basal Ganglia* (Houk, J.C. *et al.*, eds), MIT Press
23 Montague, P.R. *et al.* (1995) Bee foraging in uncertain environments using predictive Hebbian learning. *Nature* 377, 725–728
24 Egelman, D.M. *et al.* (1998) A computational role for dopamine delivery in human decision-making. *J. Cognit. Neurosci.* 10, 623–630
25 Wise, R.A. *et al.* (1978) Neuroleptic-induced anhedonia in rats: pimozide blocks reward quality of food. *Science* 201, 262–264
26 Daw, N.D. *et al.* (2002) Opponent interactions between serotonin and dopamine. *Neural Netw.* 15, 603–616
27 Dickinson, A. (1985) Actions and habits – the development of behavioral autonomy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 67–85