Neuropsychopharmacology (2007) 32, 54–66
© 2007 Nature Publishing Group All rights reserved 0893-133X/07 $30.00
www.neuropsychopharmacology.org

Linking Animal Models of Psychosis to Computational Models of Dopamine Function

Andrew J Smith[1], Ming Li[2], Suzanna Becker[3] and Shitij Kapur*[,4]

[1]Bentley Heath, Solihull, West Midlands, UK; [2]Department of Psychology, University of Nebraska-Lincoln, Lincoln, NE, USA; [3]Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, ON, Canada; [4]Centre for Addiction and Mental Health, Department of Psychiatry, University of Toronto, Toronto, ON, Canada

Psychosis is linked to dysregulation of the neuromodulator dopamine and antipsychotic drugs (APDs) work by blocking dopamine receptors. Dopamine-modulated disruption of latent inhibition (LI) and conditioned avoidance response (CAR) have served as standard animal models of psychosis and antipsychotic action, respectively. Meanwhile, the 'temporal difference' algorithm (TD) has emerged as the leading computational model of dopamine neuron firing. In this report TD is extended to include action at the level of dopamine receptors in order to explain a number of behavioral phenomena including the dose-dependent disruption of CAR by APDs, the temporal dissociation of the effects of APDs on receptors *vs* behavior, the facilitation of LI by APDs, and the disruption of LI by amphetamine. The model also predicts an APD-induced change to the latency profile of CAR—a novel prediction that is verified experimentally. The model's primary contribution is to link dopamine neuron firing, receptor manipulation, and behavior within a common formal framework that may offer insights into clinical observations.
Neuropsychopharmacology (2007) **32,** 54–66. doi:10.1038/sj.npp.1301086; published online 10 May 2006

Keywords: dopamine; temporal difference learning; psychosis; conditioned avoidance; latent inhibition; computational modeling

INTRODUCTION

Perturbations in dopamine transmission are central to a number of human illnesses including addiction, Parkinson's disease, attention-deficit hyperactivity disorder (ADHD), and schizophrenia. As a result there has been tremendous interest in understanding the psychological and behavioral functions of dopamine, and a number of overlapping hypotheses have been suggested. These hypotheses include roles for dopamine in hedonia (Wise, 1982), reward learning (Robbins and Everitt, 1996), incentive salience (Berridge and Robinson, 1998), sensorimotor function and anergia (Salamone *et al*, 1994). In parallel, mathematical models have been used to link some of these psychological constructs to the physiology of the dopamine system (Schultz, 1998; Seamans and Yang, 2004). Many of these models have concentrated on 'normal' dopaminergic conditions (McClure *et al*, 2003; Schultz *et al*, 1997), and the current aim is to explore how one such model can be extended to address the abnormal conditions encountered in schizophrenia and their treatment.

Of all the mathematical models linking dopamine, behavior, and psychology, the Temporal Difference Learning model (TD) has enjoyed particular success (Montague *et al*, 1996; Redish, 2004; Schultz *et al*, 1997). TD is a powerful formal reinforcement learning technique that has been used to solve many challenging machine learning problems. For example, the technique has been used to train computers to play backgammon to the highest human standards by associating a simulated reward with winning a game (Tesauro, 1994). At the heart of TD is a *prediction-error signal* that is zero if the environment behaves as expected, positive in response to unexpected reward, and negative following omission of expected reward. Striking parallels have been observed between the firing of dopamine neurons in monkeys and the prediction-error signal in TD simulations (Bayer and Glimcher, 2005). For example, dopamine neurons show burst firing in response to unexpected reward or unexpected predictors of reward (Romo and Schultz, 1990), an absence of firing in response to rewards that are expected (Schultz *et al*, 1992), and below baseline depressions in firing in response to omission of an expected reward (Hollerman and Schultz, 1998). Extensions of TD from neuron firing to expressed behavior of animals have been suggested (Dayan and Balleine, 2002; McClure *et al*, 2003; Montague *et al*, 1995), creating an opportunity

*Correspondence: Dr S Kapur, Centre for Addiction and Mental Health, Department of Psychiatry, University of Toronto, Toronto, ON, Canada. E-mail: andys_kathmandu@yahoo.com
Received 30 September 2005; revised 8 March 2006; accepted 9 March 2006
Online publication: 23 March 2006 at http://www.acnp.org/citations/Npp032306050606/default.pdf

to relate the TD model of dopamine function to a vast array of experimental animal data.

Dopamine plays a central role in the pathophysiology of schizophrenia, particularly with respect to the manifestation of psychosis (ie delusions, hallucinations etc.) (Abi-Dargham, 2004). Every clinically effective antipsychotic drug (APD) blocks dopamine D2 receptors (Kapur and Mamo, 2003), and dopamine-enhancing agents such as amphetamine can induce psychomimetic symptoms in otherwise normal people (Connell, 1958). Given the technical and ethical limitations of direct experiments in patients, animal models of altered dopamine function have provided a major venue for understanding both the pathophysiology and treatment of schizophrenia. For example, conditioned avoidance (CA) is a classic animal model in the study of antipsychotic drugs and their dopamine-blocking properties (Wadenberg *et al*, 2001), and one that has been used extensively as a pre-clinical screen for antipsychotic efficacy (Janssen *et al*, 1965). Meanwhile, latent inhibition (LI) is widely used in the study of selective attention in the context of reward learning. LI is disrupted not only in animals and people following induced-hyperdopaminergic states but also in patients with schizophrenia (Weiner, 2003). There is a rich history of animal and human experimentation that supports LI disruption as a plausible model of the processing deficits seen in schizophrenia (Lubow, 2005), and it is of great interest to determine the applicability of TD to altered dopamine function in both these behavioral paradigms.

The primary contribution of this paper is to demonstrate that the TD model of dopamine neuron firing can be extended to account for animal *behavior* in CA during manipulation of the dopamine D2 receptor via systemically administered haloperidol. In order to model the effects of pharmacological manipulation on behavior, we make the following assumptions that link a computational model to biology.

(a) The firing of dopamine neurons represents a TD-like prediction error signal.
(b) Burst-firing of dopamine neurons leads to phasic increases in dopamine within the synapse.
(c) The prediction error is detected by synaptic dopamine receptors, and then processed downstream of those receptors.
(d) Haloperidol blocks the effect of phasically released dopamine on the intrasynaptic dopamine D2 receptors, thereby pharmacologically decreasing the prediction error as it is perceived downsteam of the dopamine receptor.
(e) Amphetamine enhances the effect of phasically released dopamine on the intrasynaptic dopamine receptors, thereby pharmacologically increasing the prediction error as it is perceived downstream of the dopamine receptor (see Redish (2004) for a related idea).

In TD, the prediction error signal is used to update internal estimates of the reward or punishment associated with conditioned stimuli. Our approach is to model haloperidol by subtracting a dose-dependent constant from the prediction error before it is used to update the internal estimates. Conversely, amphetamine is modeled by *adding* a constant to the prediction error before it is used to update the internal estimates. Note that it is not necessary to argue that haloperidol attenuates the phasic release itself, simply that the *impact* of that release on the D2 receptor is reduced due to receptor blockade, making the prediction error appear smaller than it actually is. Following McClure *et al* (2003), we make the final assumption that the TD internal estimates of rewards are used by the animal to motivate behavior. Therefore, correlations are sought between the magnitudes of these internal estimates within the model, and the strength/probability of observing the conditioned avoidance response in animals. These assumptions are evaluated in the discussion.

Abstract computational models of such phenomena cannot be evaluated in terms of the precision with which they match any single experiment. This is because there always exist some model parameters that are biologically underconstrained and that can be hand-tuned, *ad-hoc*, for the best possible quantitative match. Therefore we seek a broad *qualitative* correspondence between experimental data and model output, but one that *generalizes* to other experiments with minimal alteration to the model. Towards this aim, in the second section of this paper, the same model of CA is applied to historical and independently collected data from LI experiments. Within the model, we make the assumption that LI can be accounted for by a pre-exposure-induced retardation in conditioning. Under this assumption, the same model used to account for effects of dopamine manipulations on CA is also shown to account for the effects of dopamine manipulations on LI. In the light of these computational analyses, the role of dopamine in these animal models is re-interpreted to reflect processes of prediction error, reward learning, and attribution of incentive salience.

DEVELOPING THE MODEL FOR CONDITIONED AVOIDANCE

Introduction

In a typical conditioned avoidance experiment, a rat is placed in a two-compartment shuttle box and presented with a neutral conditioned stimulus (CS) such as a light or tone for 10 s, immediately followed by an aversive unconditioned stimulus (US), such as a foot-shock. The animal may *escape* the US when it arrives by running from one compartment to the other. However, after several presentations of the CS-US pair, the animal typically runs during the CS and before the onset of the US, thereby *avoiding* the US altogether. It is well established that dopamine antagonists such as APDs, at non-cataleptic doses, disrupt the acquisition and expression of conditioned avoidance (Courvoisier, 1956).

No consensus has been reached regarding the behavioral or psychological processes underpinning APD-induced disruption of the avoidance response. One hypotheses suggests that APDs disrupt avoidance by creating a mild 'motor initiation' deficit (Anisman *et al*, 1982; Fibiger *et al*, 1975). An alternative hypothesis states that in addition to any effect on the motor system, APDs decrease CA by impairing reward learning and by hindering the attribution of motivational salience to the CS (Beninger, 1989a, b).

According to the 'motor initiation' deficit account, avoidance disruption might be expected to possess three characteristics: (a) manifestation that is contemporaneous with the presence of the drug; (b) dose dependency; (c) little or no evidence of disruption once the drug has left the system. This scenario, labeled 'simple motor disruption' in Figure 2 (top), is illustrated for three hypothetically increasing doses of dopamine blockade with interspersed drug-free sessions. According to the alternative 'motivational salience' hypothesis, dopamine is required for the dynamic attribution of incentive salience to the CS, and to mediate the CS's ability to motivate an avoidance response. Both models make similar predictions about overall session effects and simple dose-dependency, but the critical difference relates to trial-by-trial variation within the session. Unlike the simple motoric model, a learning-based account predicts that APD-induced disruption will gradually manifest itself as the session progresses, even though dopamine receptor blockade is stable. Furthermore, disruption in a subsequent drug-free session should be influenced by what was learned (under the influence of the drug) in the previous session. TD exemplifies the idea that dopamine mediates reward learning, and Figure 2 (middle) contrasts the TD model predictions to those of the motoric model, in the same simulated experiment.

Materials and Methods

In this section, the TD method used to generate Figure 2 (middle) is outlined, followed by a description of an animal experiment used to distinguish between Figure 2 (top) and (middle).

Modeling CA with TD. TD operates by learning to estimate the total expected future reward or punishment following each simulated stimulus. Learning is driven by a *prediction error* signal and it is this signal that is thought to correspond to dopamine neuron firing. In the case of CA, consider two simulated environment *states*: one representing the CS, and one representing the US. Additionally, in order to allow TD to capture the observed sensitivity of dopamine neurons to inter-stimulus intervals (Hollerman and Schultz, 1998), a number of internal interval or timing states are also conventionally assumed. These internal timing states are denoted here as I1, I2, etc and are shown in Figure 1. It is unclear how animals represent event timing and there is no way of knowing exactly how many of these timing states to use. Fortunately, it is usually the case that the final behavior of the model depends only quantitatively, and not qualitatively, on this decision. Each state has a fixed intrinsic *reward* associated with it, and since the US is the only intrinsically reinforcing state in this simulation, it is assigned a reward of 1, while all other states are assigned zero reward. The states and rewards are fixed at the beginning of the simulation and effectively model the animal's environment.

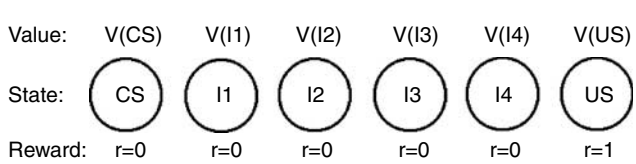

Figure 1 Representing conditioned avoidance. Each circle represents a distinct *state* of the environment, and is associated with a Value (*V*) and a reward (*r*). *r* is assumed to be provided by the environment and will be zero for all neutral states. As the US is the only intrinsically rewarding state in this task, this is the only state to have a non-zero reward. The Value, *V*, is the model's estimate of the total future reward expected to follow that state and must be learned through experience.

Also associated with each state is a *value*. The Value (henceforth capitalized to distinguish the special meaning) is the agent's internal estimate of the total future reward expected to follow that state. Initially, all these Values are zero (the agent starts off naive), but will be adapted through experience. In this particular example, we expect all the Values to eventually adapt towards 1 because $r(\mathrm{US}) = 1$, and this comprises the only future reward following the CS. However, in the general case, there could be many different states following the 'CS', each encountered with different probabilities and each with their own intrinsic reward. In such cases, the Value of the CS is adapted during learning to reflect the expected total reward following that CS—that is, a sum of all future consequences. If the states and rewards are a model of the animal's environment, the Values are a highly abstract model of the information learned by the animal and subsequently used to motivate behavior. The Values can be interpreted as the agent's representation of *incentive salience* (McClure *et al*, 2003), and are used within machine learning to drive behavior. For example, the greater the Value of a state, the more future reward is expected and the more incentive exists for acquiring or avoiding that state.

To make the problem easier to model, each trial is broken up into discrete time points, with each new time point corresponding to a new state. For example, at time $t = 1$ the current state is 'CS', and at $t = 2$ the current state is 'I1', etc. For the current time point, t, a prediction error is generated which can be loosely paraphrased as 'actual future reward−expected future reward' and reflects the error in the Value of the current state:

$$\delta(t) = (r(t) + \gamma V(t+1)) - V(t) \quad (1)$$

where $V(t)$ is the Value of the current state, $r(t)$ is the intrinsic reward associated with the current state, $V(t+1)$ is the Value of the next environmental state, and $\gamma \leqslant 1$ is an experimenter-selected *discount factor* that controls the degree to which future rewards are discounted over immediate rewards (eg a dollar tomorrow is worth slightly less than a dollar today, etc.). The prediction error of Equation (1) is then used to improve the Value of the current state:

$$V(t) = V(t) + \alpha\delta(t) \quad (2)$$

where $0 \leqslant \alpha \leqslant 1$ is a learning rate which we assume to be controlled by biological parameters outside the scope of the model. The Values are gradually adapted over a number of trials by invoking Equations (1 and 2) at each time point in each trial. The theory behind Equations (1 and 2) was originally developed to tackle a set of complex engineering problems, and it is therefore striking that Equation (1) has been shown to resemble the firing patterns of actual dopamine neurons. In order to compare the output of the TD model to animal behavior, McClure *et al* (2003) have

suggested that V can be interpreted as an abstract measure of incentive salience (Berridge and Robinson, 1998) or motivation. Although over simplifying animal behavior, the basic principle that an animal is motivated by the future reward predicted by a stimulus is both intuitive and convenient. We therefore assume that $V(t)$ can be interpreted as motivation to produce an avoidance response in the current state. This relationship is formalized in the simplest possible way by defining the probability of producing an avoidance response in the current state as: $p(t) = V(t)$. As the first avoidance response of a trial will end that trial, the overall probability of observing an avoidance response at time t is:

$$\hat{p}(t) = p(t) \times \left(1 - \sum_{i=1}^{t-1} \hat{p}(i)\right) \tag{3}$$

Equation (3) simply states that $\hat{p}(t)$ equals $p(t)$ multiplied by the probability of not having already produced an avoidance response in that trial (which would have ended

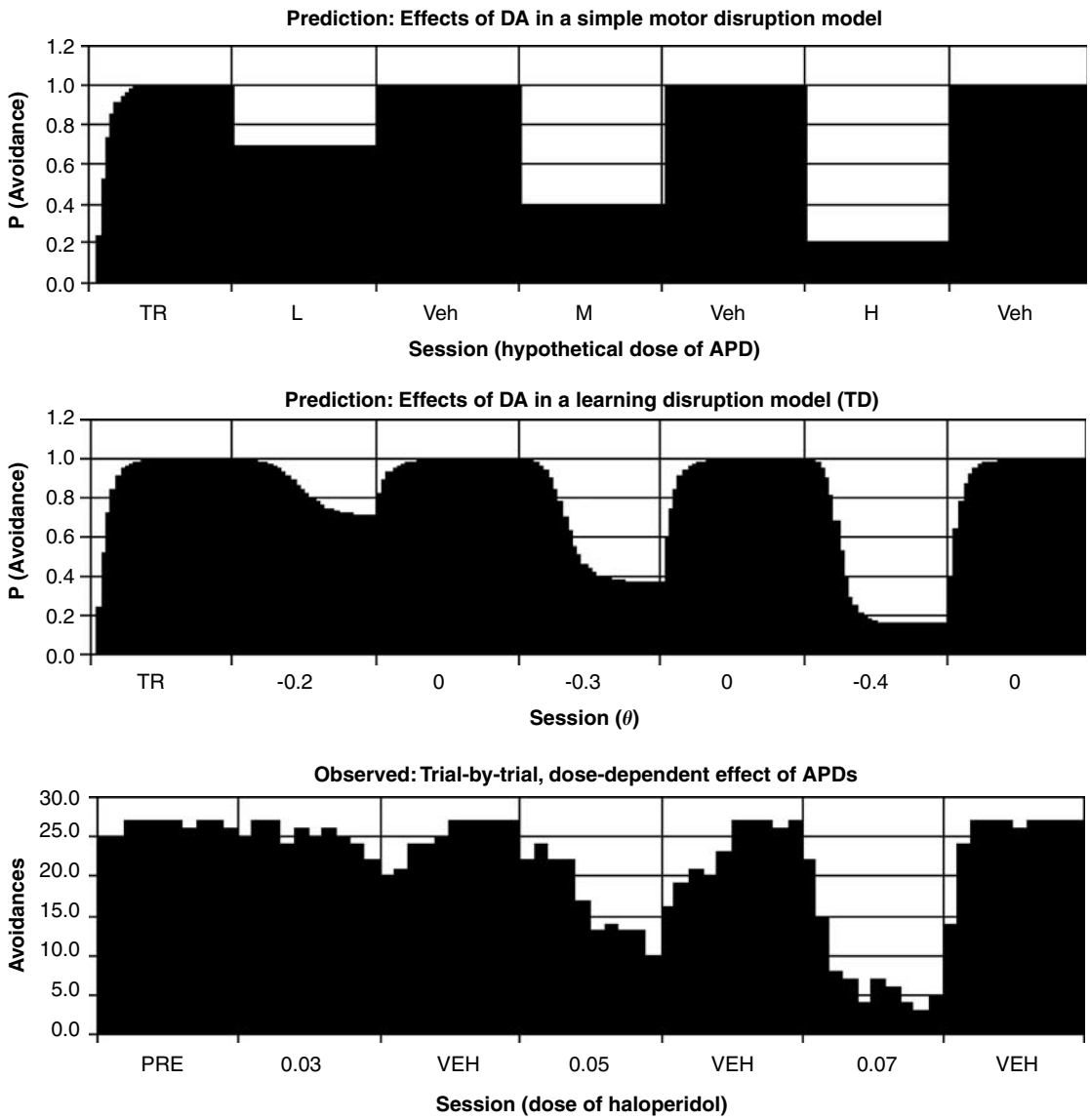

Figure 2 Simulated and actual profiles of avoidance disruption over seven consecutive sessions. *Top*: A simple hypothesis in which avoidance disruption is proportional to receptor blockade The probability of producing an avoidance response is plotted over a hypothetical experiment involving seven consecutive sessions of 30 trials per session. Each hypothetical session involves a different, stable, level of dopamine receptor blockade achieved with either low (L), medium (M), or high (H) doses of APD or vehicle (Veh). This hypothetical pattern characterizes immediate or *performance-like* effects. The session labeled 'TR' denotes initial drug-free training. *Middle*: In contrast, the behavior of the TD model is displayed, again in terms of the probability of producing an avoidance response. Each session is modeled (and labeled) using a different value of θ in Equation (4). Note that although θ is stable throughout each session, avoidance continues to change within each session. 'TR' refers to drug-free training in which $\theta = 0$. The consequence of simulated dopamine blockade ($\theta < 0$) is a pharmacologically induced reduction of the prediction error, δ_{rec}, and an incremental dampening of the Values below their true equilibrium. If a drug-free session is then administered ($\theta = 0$), this discrepancy is gradually corrected and the full Values are restored. *Bottom*: Total number of avoidances in a two-way CA animal experiment (see text) comprising seven consecutive sessions, each separated by at least 48 h. Each session consisted of 30 trials, shown in blocks of three. The dose of APD (haloperidol) administered 1-h before each session is shown on the *x*-axis. 'VEH' refers to vehicle. The performance on the last (of 11) drug-free training day is also shown and labeled 'PRE'.

the trial). Now, $\hat{p}(t)$ should be proportional to the number of rats producing an avoidance response during the interval of time represented by the current state. The overall probability of producing an avoidance response can be calculated by $\hat{p}(CS)+\hat{p}(I1)+\hat{p}(I2)+\hat{p}(I3)+\hat{p}(I4)$, where we have replaced the variable, t, with the corresponding state.

Note that in some trials the model will produce an avoidance response early, thus ending the trial and avoiding the shock. In these cases, the Values of the later interval states and US state are not updated because the simulation, like the animal, does not experience them. After sufficient trials and applications of Equation (2), $V(CS)\approx V(I1)\approx V(I2)\approx V(I3)\approx V(I4)\approx V(\mathrm{US})\approx 1$ (if $\gamma\approx 1$). At this point every state accurately predicts the future reward of the US, and there will be no residual prediction errors and no more learning. Figure 2 (middle, labeled 'TR') shows how the overall probability of producing an avoidance response increases from 0 to 1 during thirty trials (simulating initial training in CA).

The simplest approach to simulating dopamine receptor manipulation is to add some constant, θ, to all prediction errors:

$$\delta_{\mathrm{rec}}(t)=\delta(t)+\theta \tag{4}$$

This distinguishes between the phasic dopamine response, $\delta(t)$, which is assumed to be proportional to the amount of dopamine released into the synapse, and the *effect* of that release downstream of the dopamine receptor, $\delta_{\mathrm{rec}}(t)$. Normally the two will be the same ($\theta=0$), but when dopamine receptors are blocked within the synapse, $\theta<0$ can be used to simulate the reduced impact of normal dopamine release at those receptors. From now on, $\delta_{\mathrm{rec}}(t)$ is used in place of $\delta(t)$ in Equation (2) reflecting the assumption that modification of $V(t)$ happens downstream of the dopamine receptor.

To summarize, each time-step in each trial is simulated within the model by:

(1) Generation of prediction error (assumed to correspond to phasic dopamine response).
(2) Haloperidol dose-dependent constant being subtracted from prediction error. The new quantity is interpreted as the drug-reduced *impact* of the phasic release on the D2 receptor within the synapse.
(3) The drug-modulated prediction error is used to update the Values. Adaptation of these representations is assumed to occur downstream of the blocked D2-receptor.
(4) The probability of producing an avoidance response is generated based on the current Value for the current state.

Figure 2 (middle) shows how the probability of producing an avoidance response using TD varies from trial to trial in the same simulated experiment as Figure 2 (top). Simulated blocking of the D2 receptor leads to a gradual reduction in the Values (and therefore avoidance response). Subsequent simulation of drug-free trials leads to a gradual restoration of the Values. For more details on the TD algorithm and the parameters used, see Supplementary Materials.

Animal experiment. In pharmacological studies of APDs in CA, aggregate session data are commonly reported. However, in order to determine which of the scenarios of Figure 2 (top and middle) is more accurate, trial-by-trial data are required for interleaved drug/drug-free sessions. Therefore, a tailored CA experiment was performed.

Twenty-four rats were first divided into three groups, each group being trained and tested with a different CS–US interval: 6 s ($n=8$), 12 s ($n=7$), and 24 s ($n=9$). Once allocated, each rat was only ever exposed to trials involving the relevant CS–US interval. Three CS–US intervals were used because the number of interval states in the model is underconstrained and it was therefore convenient to have experimental data pertaining to multiple intervals.

The training phase consisted of 11 sessions, separated by at least 2 days. Detailed description of the apparatus and the procedure can be found in Li *et al* (2004). Briefly, for each session, each subject was placed in a two-compartment shuttle box. A trial started by presenting a white noise (CS, 74 dB, 6, 12 or 24 s) followed by a scrambled footshock (10 s, 0.8 mA). The subject could avoid the shock by shuttling from one compartment to another during the CS. Shuttling during the shock turned the shock off. Each session consisted of 30 trials separated by a random inter-trial interval (30–60 s). By the end of the training phase all rats showed avoidance performance above 75% criterion, except one rat in the 6 s group which was dropped from the experiment. Resultant mean avoidance was above 90% for all three groups. The only addition to the previously published procedure/apparatus was a barrier between the two compartments (4 cm high), so the rats had to jump from one compartment to the other.

After the last day of training, the drug testing started. Exactly the same procedure was employed during testing, except that 1 h before each testing session one of three doses of haloperidol was administered: 0.03, 0.05, and 0.07 mg/kg haloperidol. At least 4 days were allowed to elapse between each drug-session, and one vehicle retraining session was given during that interval to maintain a high level of avoidance responding. This experiment was a within-subject design, with all groups of rats tested in the same order: 0.03 mg/kg, Veh, 0.05 mg/kg, Veh, 0.07 mg/kg, Veh. The sequence therefore reflected the schedule of Figure 2 (top and middle). The 1-h interval between drug administration and testing was to ensure stable receptor blockade during testing. Note that three doses, 0.03, 0.05, 0.07 mg/kg of Haloperidol were modelled by $\theta=-0.2$, $\theta=-0.3$, $\theta=-0.4$, respectively, in Equation 4. These model values were selected for the best match to the data. The vehicle dose was modelled by $\theta=0$.

Results

For all three groups, a dose-dependent, within-session decline in avoidance responding during each drug session was observed, along with a gradual recovery of the response in the subsequent drug-free session. This within-session decline under haloperidol and within-session recovery under vehicle is shown in Figure 2 (bottom) for the 24 s group (the other two groups were almost identical). The trial-by-trial data show that under a new dopaminergic state (be it a transition to dopamine blockade, or to vehicle), the animals' behavior supports the TD hypothesis of a gradual learning curve. It is important to note that within the model,

this learning effect is not a result of changes to the hedonic impact of the US (r(US) remains unaffected by simulated APD), but rather to the way in which the reward is processed in terms of the attribution of Value to conditioned stimuli via the prediction error.

The behavior of the TD model is defined not just for each individual trial, but also for each interval state within each trial (c.f. Figure 1). Figure 3 (black bars, top and bottom) shows how, under APDs, the avoidance response initiation becomes increasingly delayed as the session progresses. This is reflected both in the experimental data and the TD simulation. Within the TD model, this delayed initiation may be understood by considering that alterations to the prediction error signal will have an impact at every state. A cardinal feature of TD is that the Value at any given state '*X*' is dependent on the Values at all other states that intervene between '*X*' and the reward. Therefore, the more intervening states, the greater the number of prediction errors involved in adapting the Value of '*X*'. Modifications to these prediction errors will therefore have an accumulative effect the further '*X*' is from the actual reward. In short, the Values (white bars in Figure 3 top), of distal predictors will be more affected by simulated APDs leading to a change in the probability distribution of the avoidance response (black bars in Figure 3 top).

Discussion

Disruption of CA has traditionally been attributed either to motor initiation deficits (Anisman *et al*, 1982; Fibiger *et al*, 1975) or to disrupted attribution of motivational salience to the CS (Beninger, 1989a, b). The TD model interprets dopamine as mediating a prediction error, the manipulation of which leads to incremental changes in the attribution of incentive salience to the conditioned stimulus. The model is able to account for the gradual, dose-dependent, reduction in avoidance responding during each session under haloperidol. The reduction in avoidance responding in the model is due to an incremental reduction of the incentive salience of the CS, which is in turn due to a drug-induced reduction to the prediction error—a reduction that is constant (dose-dependent) throughout the session. The model also accounts for the observation that avoidance responding at the beginning of each session is comparable to the avoidance responding at the end of the previous session, irrespective of the drug treatment in either session. This across-session effect in the model is due to the fact that the incentive salience of the CS only changes in response to CS–US exposure. In the model, manipulating dopamine has no effect whatsoever, unless the CS–US stimuli are also present. Therefore, between sessions, no change in incentive salience is observed, even though dopamine receptor manipulation changes. The model also accounts for the observed response-initiation effects (see Figure 3). Within the model, this effect is due to the fact that manipulation of the prediction error has a greater effect on states that are distant from the US.

The ability of TD to account for these behavioural observations arising out of pharmacological manipulation of the dopamine receptor builds on the already large literature that uses exactly the same TD model to account for the phasic firing of dopamine neurons. The basic hypothesis is that the US causes a phasic dopamine response that conveys a prediction error, which in turn increases the Value (incentive salience) of the CS. A full discussion of this literature is out of scope of the current account (but see Schultz (1998) for a review). By pharmacological manipulation of the perceived phasic

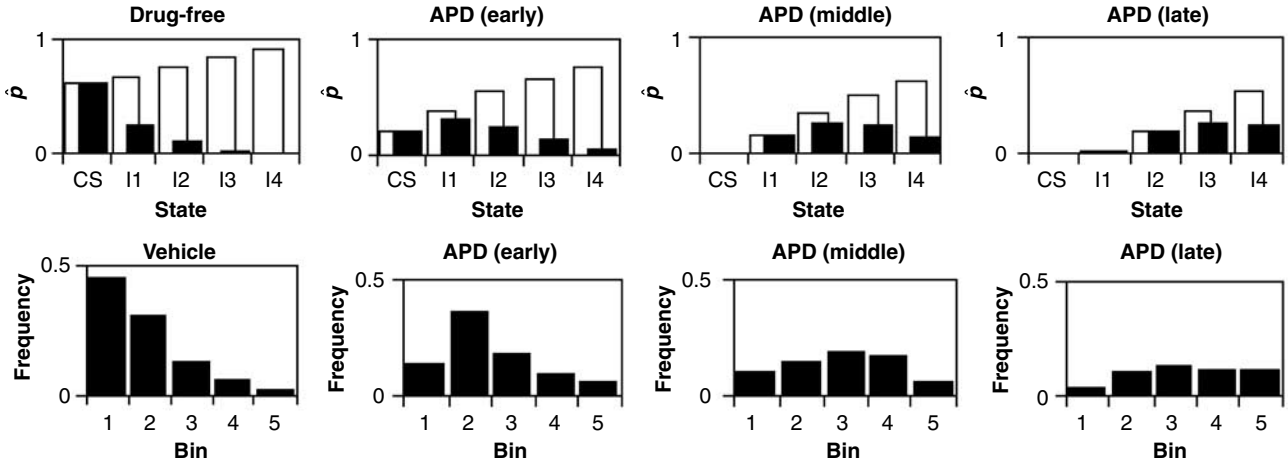

Figure 3 TD predicts within-session increases in avoidance response latency. Each graph shows the duration of the CS along the *x*-axis, divided either into the states of Figure 1 (*top*, TD), or five temporal bins (*bottom*, experimental data). *Top*: The Value of each of the five 'avoidance' states are denoted by the white bars, showing that simulated dopamine blockade has a greater cumulative impact on the 'early' predictors of the US. The black bars denote the corresponding probability of observing a simulated avoidance response in that state (see Equation (3)). The leftmost graph shows the behavior of the model in a fully trained, drug-free trial. The right three graphs show the behavior of the model at three stages (early = trial 4, middle = trial 7, late = trial 9) in the simulated 0.05 mg/kg session (ie $\theta = -0.3$). *Bottom*: Experimental data. Each of the four plots show the proportion of avoidance responses falling within each of five bins during the CS. The 12 s group is shown here, so each bin corresponds to an interval of 12/5 = 2.4 s. The leftmost plot shows the performance on the final day of training and is typical of baseline, drug-free behavior. In the drug-free case, the majority of responses occur in the first bin. The three plots to the right show how the avoidance profile changes in the beginning (trials 1–10), middle (trials 11–20) and end (trials 21–30) of the session, even though dopamine receptor blockade is presumed to be approximately constant. Notably, response initiation becomes increasingly delayed. This effect was similar for the 6 and 24 s groups.

response, the current model allows the prediction error to be modified and thereby also the incentive salience of the CS. The effects of pharmacological dopamine manipulation on the firing of dopamine neurons, and the implications for the TD model of dopamine neuron firing, remain untested. Incidentally, within the TD model, the Value associated with the CS (for example the black bars in Figure 3 (top)) predicts the phasic firing of dopamine neurons under these behavioural/pharmacological conditions. These predictions await testing.

GENERALIZING THE MODEL TO LI DISRUPTION

Introduction

Latent inhibition refers to a subject's increased difficulty to form a new association between a stimulus and a reward due to prior exposure of that stimulus without consequence (Lubow and Moore, 1959). The behavioral hallmark of LI is retarded conditioning of the pre-exposed stimulus, and is thought to reflect the ability of organisms to ignore irrelevant stimuli (Lubow, 1989). In two pivotal studies, Solomon *et al* (1981) and Weiner *et al* (1981) both reported the disruption of LI (failure of pre-exposure to retard subsequent conditioning) in animals treated with amphetamine (a dopamine-enhancing agent). Links with the pharmacology of psychosis are reinforced by findings that APDs *reverse* the amphetamine-induced LI deficit in animals (reviewed in Moser *et al*, 2000) and *enhance* the standard LI effect in normal animals (Weiner *et al*, 1987). Based on this pharmacological connection, LI-disruption was suggested as a potential model of the selective attention deficit found in schizophrenia. Subsequently, LI has been found to be disrupted in acute psychotic schizophrenic patients (Baruch *et al*, 1988), and in normal humans treated with amphetamine (Gray *et al*, 1992). These and other evidence support the validity of LI-disruption as a model of processing deficits in acute schizophrenia.

Materials and Method

A common LI procedure involves a *pre-exposure* phase during which a neutral stimulus is repeatedly presented without consequence, followed by a conditioning/testing phase during which the same stimulus (now being used as a CS) is followed by a US. LI is indexed by the impaired acquisition of the conditioned response in pre-exposed animals compared to non pre-exposed animals. For example, Weiner *et al* (1988) pre-exposed one group of rats (PE) to a 5 s tone, without consequence. A second group (NPE) was not pre-exposed to the tone. Both groups were then trained in a two-way conditioned avoidance experiment for 60 trials. Each group was subdivided before the experiment began. One sub-group (AMPH) was administered with amphetamine before CA training, and one subgroup (SAL) was administered with saline. This yielded four groups, AMPH(PE), AMPH(NPE), SAL(PE), SAL(NPE), and permitted an investigation into the interactions of amphetamine and pre-exposure with respect to LI (See Supplementary Material for details of the experimental method). Figure 4 (right) summarizes the results of the experiment, confirming both the LI affect (SAL(PE) *vs* SAL(NPE)) and the disruption of LI in the amphetamine-treated group (SAL(NPE) *vs* AMPH(PE)). Note that LI can be assessed and modeled in a variety of behavioral contexts, but CA is considered here for consistency with the TD model used earlier.

The basic TD model does not directly account for LI, since supplying the CS on its own during pre-exposure yields no prediction error, no changes to incentive salience, and no changes to how incentive salience is subsequently assigned to the CS. However, under the assumption that pre-exposure retards the conditioning rate of the CS, we demonstrate that the same TD model used to account for the haloperidol-induced disruptions to CA can be used to account for dopaminergic interactions with LI.

Mackintosh (1974) suggested that the pre-exposure phase decreases the associativity of the CS, and we therefore attach an 'associativity' parameter, ϕ, to the CS which controls the rate of conditioning of that CS (Note the difference between θ and ϕ). We abstract over the debated and potentially detailed mechanisms by which pre-exposure may reduce associativity by simply assuming that a non pre-exposed CS has $\phi = 1$ and a pre-exposed CS has $0 \leqslant \phi \leqslant 1$ which is constant following pre-exposure. Equation (2) now becomes:

$$V(t) = V(t) + \alpha\phi\delta_{\mathrm{rec}}(t) \tag{5}$$

effectively reducing the learning rate of a pre-exposed stimulus. We could equally assume that pre-exposure acts directly on the learning rate parameter. However the learning rate parameter is intended to capture all those miscellaneous and unknown biological factors that contribute to the rate at which an association is learned, while ϕ is intended to specifically capture the effect of interest—namely pre-exposure. The two approaches are mathematically equivalent. It is important to stress that the pre-exposure phase itself will not involve any iterations of the model's learning equations. TD effectively does nothing during pre-exposure. The introduction of ϕ assumes some unspecified mechanism has already retarded the associativity of the CS, and thereby also its ability to condition.

Figure 4 (right) confirms that the assumption of Equation (5) does indeed yield LI under simulated pre-exposure ($\phi = 0.4$). The same model used to account for haloperidol-induced disruption of CA can now be used to account for the interactions between dopamine manipulation and LI. These interactions can be simulated with TD using exactly the same mechanism that was used previously for CA. For example, haloperidol is again simulated by $\theta < 0$ (not to be confused with associativity, ϕ!), while the dopamine-enhancing properties of amphetamine are now simulated by $\theta > 0$. Modeling amphetamine by adding a positive constant to the prediction error is intended to capture the enhanced impact of the normal phasic response as a result of blocking re-uptake within the synapse. However, enhancing the prediction error in this way can ultimately lead to Values greater than 1. Since the probability of producing an avoidance response is linked to the Values, and a probability cannot exceed one, we add the constraint that $p(t) = V(t)$ or 1, whichever is least. As in CA, θ is assumed not to affect directly the amount of dopamine actually released, but rather the *impact* of that release (although see discussion). For clarification, note that θ acts in the 'post-

receptor' prediction-error of Equation (4) and not in the 'phasic response' prediction error of Equation (1).

Results

Figure 4 (right) demonstrates how the model exhibits LI-disruption under simulated amphetamine. Within the model, LI occurs because conditioning is retarded by $\phi < 1$ (simulated pre-exposure), while amphetamine-induced LI disruption occurs because conditioning is enhanced by inflated prediction errors ($\theta > 0$).

The model also makes predictions pertaining to the effect of administering APDs in LI in the absence of amphetamine treatment. In short, simulating APDs with $\theta < 0$ will act in the same direction as simulating pre-exposure with $\phi < 1$. In other words simulated APDs will exaggerate retardation in conditioning brought about by pre-exposure and will therefore facilitate LI. Figure 5 (right) shows the model's predictions under these conditions. Figure 5 (left) shows experimental data from a similar LI experiment to that described above, except that an APD (haloperidol) rather than amphetamine was administered during both pre-exposure and subsequent avoidance conditioning (Weiner *et al*, 1987). Both the experiment and the simulation show facilitation of LI under APD.

Finally, APDs possess the ability to reverse amphetamine-induced disruption of LI (reviewed in (Moser *et al*, 2000)). Within the model, this effect is trivial as amphetamine is modeled by a positive value of θ, and APDs by a negative value. The two would therefore cancel each other if simulated simultaneously.

Discussion

The basic TD model does not directly account for LI, since supplying the CS on its own during pre-exposure yields no

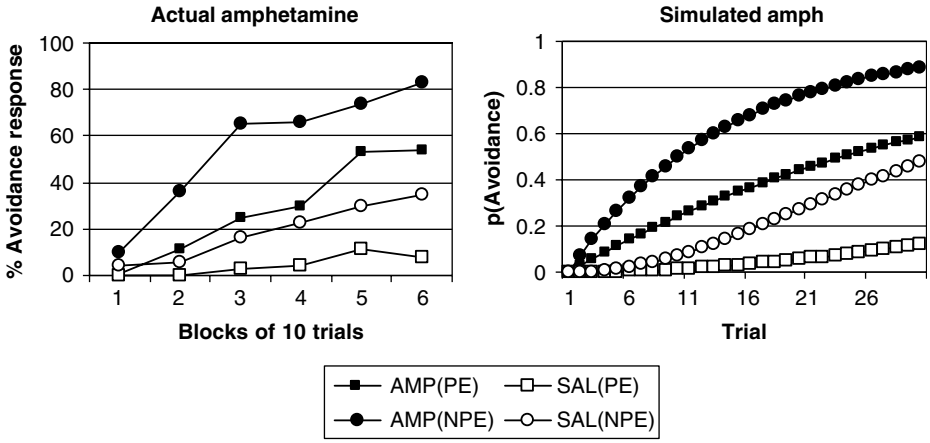

Figure 4 The LI effect, and its disruption with amphetamine, shown experimentally and within the TD model. *Left*: Acquisition rates for the four groups of rats of interest. Pre-exposed (PE) and non pre-exposed (NPE) groups were subdivided into those treated with acute doses of amphetamine (AMP) and placebo (SAL) before conditioning. Acute amphetamine is shown to disrupt LI. Adapted from Weiner *et al* (1988) Figure 2. *Right*: TD simulation. Both amphetamine treatments are simulated by setting $\theta = 0.3$ and pre-exposure is simulated by setting $\phi = 0.4$. The learning rate was empirically selected ($\alpha = 0.05$) so that the model produced an acquisition rate for the baseline group (SAL(NPE)) that was comparable to the animal experiment. The scales of the axes are largely irrelevant in models of this kind, because yoking the free parameters to biological constants is prohibitively difficult. However, the robust *qualitative* observation is that simulated amphetamine is capable of disrupting simulated LI.

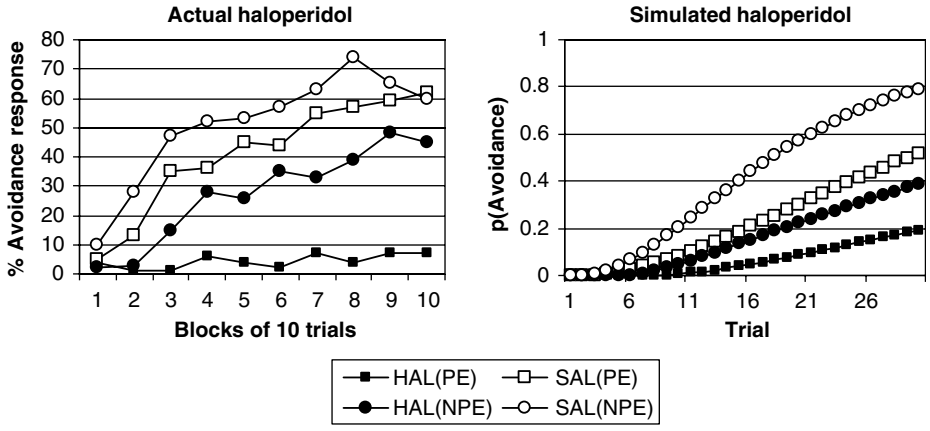

Figure 5 The LI effect, and its enhancement under the APD haloperidol, shown experimentally (left) and within the TD model (right). The experimental results are adapted from (Weiner *et al*, 1987). The model results are achieved by simulating pre-exposure with $\phi = 0.6$ and haloperidol with $\theta = -0.2$. The learning rate was arbitrarily selected ($\alpha = 0.09$) so that the model produced a quantitatively similar acquisition rate for the baseline group (SAL(NPE)).

prediction error, no changes to incentive salience, and no changes to how incentive salience is subsequently assigned to the CS. However, under the assumption that LI is due to a pre-exposure induced retardation of conditioning (modeled by a constant $\phi < 1$), the TD model can readily be made to exhibit LI. Then the same set of learning equations used to model CA-disruption can also be used to reproduce the experimentally observed interactions between dopamine manipulation and LI. The model's explanation of these observations is that manipulation of the dopamine receptor alters the dopamine-mediated 'prediction error' as it is perceived downstream of the receptor, leading to incremental changes in the attribution of incentive salience to the conditioned stimulus. In summary, simulating dopamine enhancement increases conditioning rates and asymptotes of incentive-salience while dopamine blockade has the reverse effect. These changes interact with the assumed retardation in conditioning induced by pre-exposure.

If the TD model of haloperidol-induced disruption of CA is valid, then it would be expected that the CA model could also be used to account for dopamine-LI interactions under a minimal number of additional assumptions. Although this has been demonstrated, it is recognized that other hypotheses of dopamine function may be equally able to achieve this generalization between CA and LI. In this sense, a parsimonious generalization from CA to LI is encouraging but does not provide proof of validity.

The current model can be compared with a number of existing accounts of LI. The *selective attention* hypothesis (Lubow, 1997; Mackintosh, 1975) suggests that amphetamine-induced LI disruption reflects the inability of the animal to ignore irrelevant stimuli (Solomon *et al*, 1981). Weiner's *behavioral switching hypothesis* (Weiner, 1990), on the other hand, suggests that LI disruption reflects enhanced switching from responding according to CS-no US association to responding according to CS–US association. A third alternative suggests that LI deficit is due to an increase of the salience of the reinforcer itself under amphetamine (Killcross *et al*, 1994). While the current TD model is not in a position to differentiate between these hypotheses, it does formally demonstrate the ability of simulated dopamine-reinforcer interactions to address dopaminergic disruption of both CA and LI via the constructs of prediction error and incentive salience.

DISCUSSION

Limitations

The aim has been to extend the TD model of dopamine neuron firing to address behavior and pharmacology in two animal models used in schizophrenia research. In accordance with TD theory, the heart of the model comprises Values that estimate the expected future reward following each stimulus or environment state, and a prediction error signal that updates these Values based on the difference between what is expected and what actually happens. The link between the model and biology pivots on a number of core assumptions:

(a) The firing of dopamine neurons represents a TD-like prediction error signal.
(b) Burst-firing of dopamine neurons leads to phasic increases in dopamine within the synapse.
(c) Haloperidol blocks the effect of phasically released dopamine on the intrasynaptic dopamine D2 receptors, thereby pharmacologically decreasing the prediction error as it is perceived downsteam of the dopamine receptor.
(d) Amphetamine enhances the effect of phasically released dopamine on the intrasynaptic dopamine receptors, thereby pharmacologically increasing the prediction error as it is perceived downstream of the dopamine receptor.
(e) The learned Values correspond to incentive salience or 'wanting' (Berridge and Robinson, 1998), as suggested by McClure *et al* (2003), and can be interpreted as driving motivation.

With respect to (a), the link between TD and dopamine neuron firing remains an open discussion. Although many experimental data agree with the TD model (Schultz *et al*, 1997), competing hypotheses have been suggested. These include the notions that the phasic dopamine response represents part of a behavioral switching mechanism (Redgrave *et al*, 1999), or part of a broader salience-processing mechanism (Horvitz, 2000).

With respect to (b), although there are no direct methods for measuring intra-synaptic dopamine levels with sub-second resolution (Robinson *et al*, 2003), fast-scan cyclic voltammetry studies provide strong evidence that both direct electrical stimulation of dopamine neurons and behaviorally relevant events (eg natural rewards) lead to phasic overflow of dopamine from the synapse (Garris *et al*, 1997; Phillips *et al*, 2003; Rebec *et al*, 1997; Roitman *et al*, 2004). However, it is worth noting that dopamine neuron firing and dopamine release may in fact be doubly dissociable (Garris *et al*, 1999; Grace, 1991).

With respect to (c) and (d), the doses of haloperidol used in this study have previously been shown to produce significant levels of D2-receptor occupancy (Kapur *et al*, 2003). While it is well demonstrated that acute amphetamine increases extracellular dopamine levels as measured by microdialysis (Segal and Kuczenski, 1992), its effects on intrasynaptic phasic release have not been directly demonstrated. One of the main effects of amphetamine is its blockade of the dopamine transporter, the major mechanism of quenching dopamine action in the synapse, and (Grace, 1995) has suggested that an acute dose of amphetamine leads to an augmentation of intrasynpatic phasic signalling even before extrasynaptic increases are observed (as measured by microdialysis). In the case of the LI experiment where the animals receive amphetamine twice (once at pre-exposure and once at conditioning), Joseph *et al* (2000) have shown that amphetamine enhances the impulse-dependent phasic dopamine release. Either of these mechanisms could provide the proposed synaptic enhancement of the prediction error signal. Also within the context of LI, Feldon *et al* (1991) showed that only amphetamine that causes an increase in phasic signaling, and not dopamine agonists which cause impulse-independent sustained stimulation of dopamine receptors, disrupts LI. Thus, while direct measurements of intrasynaptic dopamine are not as yet possible, the assumption that the

phasic dopamine response carries a prediction error signal that can be pharmacologically altered within the synapse seems plausible.

With respect to (e), the representations learned by animals clearly comprise more than just the attachment of future reward values to CSs (Cardinal *et al*, 2002), and indeed it has been suggested that TD offers only a partial model of reward processing in animals (Dayan and Balleine, 2002). As with many existing variants of the TD model, the tonic *vs* phasic distinction, along with extra- *vs* intra-synaptic effects, varied sub-types (D1-like *vs* D2-like etc.), anatomical distribution (striatal, limbic and cortical), and complex intracellular signaling of dopamine receptors, have been largely ignored in the current work. For example, receptor manipulation by haloperidol or amphetamine is likely to indirectly alter the quantity of dopamine released via auto-receptors and other feedback mechanisms. These and other mechanisms have not been explicitly modeled. However, in return for these abstractions over biological detail, a potentially useful parsimony has been achieved.

Implications for Psychosis

The interest in LI and CA is fuelled by their putative ability to model aspects of schizophrenia. Therefore, if a computational model such as TD can provide a framework to interpret pharmacology and behavior in these animal models, it may also have implications for psychosis in humans. While it has been known for some time that a hyperdopaminergic state leads to psychosis in schizophrenia, and while all known APDs block dopamine, there has been no convenient model to relate neurochemical findings to clinical observations. Two recent clinical developments demonstrate the current model's potential ability to bridge this gap.

First, building on the psychological notion of dopamine as a mediator of incentive salience (Berridge and Robinson, 1998), it has been suggested (Kapur, 2003) that psychosis, especially delusions, can be viewed as the result of a chaotically hyperdopaminergic system leading in turn to the assignment of aberrant levels of incentive salience to stimuli. While previous sections have focused on the impact of exaggerated incentive salience on the *probability* of producing a response in a behavioral task, here psychosis is linked to excessive and aberrant incentive salience itself. The crucial difference is that while probabilities of producing a response cannot exceed 1 (by definition), incentive salience, or $V(CS)$, is unbounded. Figure 6 contrasts the attribution of incentive salience under a normal dopamine condition with that under a chaotically hyperactive dopamine system. The hyperactive case yields an aberrantly high $V(CS)$ which, under the assumption of $V(CS) =$ incentive salience, is speculated to underpin the formation of delusions.

A second recent clinical observation relates to the time course of antipsychotic action. Contrary to the long-held presumption that there is a 'delay in onset' between dopamine-blockade and antispychotic effect (Grace *et al*, 1997) it has recently been shown that there is no such delay. Rather, the onset of antipsychotic action of APDs is almost immediate and shows a gradually increasing effect that asymptotes after a number of weeks (Agid *et al*, 2003). TD simulation of APD treatment reduces the post-receptor effect of dopamine signaling, leading to an early onset, gradual attenuation of attributed salience (c.f. Figure 2, middle and bottom). Thus, the model provides an account of two recent clinical proposals: the relationship of psychosis to dopamine via the construct of incentive salience, and the empirically observed pattern of

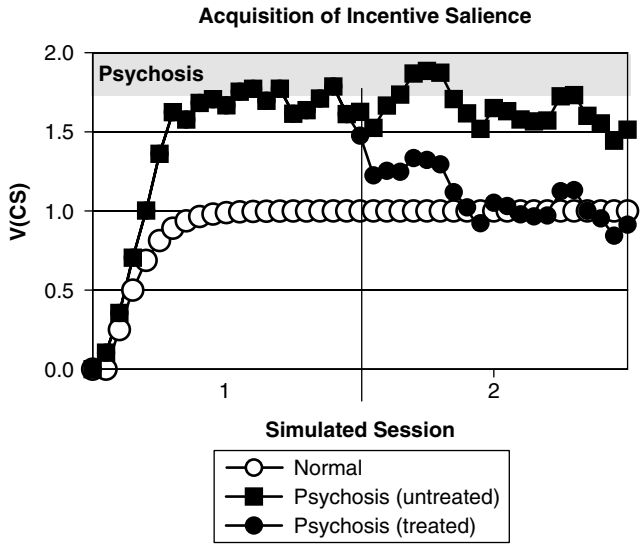

Figure 6 Implications of the model for psychosis and its treatment by dopamine-blocking drugs. The simplest version of the model is considered with just two states – CS and US. The acquisition of the Value (incentive salience) of the CS is shown over a number of trials divided into two sessions. $\alpha = 0.5$; $r(US) = 1$ throughout. The 'Normal' curve shows normal acquisition of incentive salience which asymptotes at $r(US)$. The 'Psychosis untreated' curve shows the effect of simulating a chaotic hyperdopaminergic state by randomly selecting θ from the range [0, 0.6]. $V(CS)$ is pushed above the normal asymptote of 1 yielding super-salience. An arbitrary threshold is shown at $V(CS) = 1.75$ illustrating a hypothetical clinical manifestation of psychosis. The aberrant salience can be re-learned over time by the simulated administration of APDs. This is shown in the 'Psychosis treated' curve of session 2, which is generated in the same way as the 'untreated' curve except that θ is reduced by 0.3 (ie average $\theta = 0$). One drawback of this account is that it predicts super-salience of *all* ideas, and does not explain the selectivity of delusions for some environmental features over others.

dissociation of stable receptor blockade and antipsychotic action.

In conclusion, TD provides a potentially useful computational tool for re-interpreting animal models of acute schizophrenia, and its treatment. In this respect, the model demonstrates potential for unifying electrophysiological, pharmacological, behavioral and psychological observations. However, caution is warranted by the extent to which each of these levels of description is necessarily abstracted over. The current results add to the emerging and exciting potential of TD to address human disorders in which dopamine dysfunction is implicated.

REFERENCES

Abi-Dargham A (2004). Do we still believe in the dopamine hypothesis? New data bring new evidence. *Int J Neuropsychopharmacol* **7**(Suppl 1): S1–S5.

Agid O, Kapur S, Arenovich T, Zipursky RB (2003). Delayed-onset hypothesis of antipsychotic action: a hypothesis tested and rejected. *Arch Gen Psychiatry* **60**: 1228–1234.

Anisman H, Irwin J, Zacharko RM, Tombaugh TN (1982). Effects of dopamine receptor blockade on avoidance performance: assessment of effects on cue-shock and response-outcome associations. *Behav Neural Biol* **36**: 280–290.

Baruch I, Hemsley DR, Gray JA (1988). Differential performance of acute and chronic schizophrenics in a latent inhibition task. *J Nerv Ment Dis* **176**: 598–606.

Bayer HM, Glimcher PW (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**: 129–141.

Beninger RJ (1989a). The role of serotonin and dopamine in learning to avoid aversive stimuli. In: Archer T, Nilsson L (eds). *Aversion, Avoidance, and Anxiety: Perspective on Aversively Motivated Behavior*. Lawrence Erlbaum Associates: Hillsdale, New Jersey. pp 265–284.

Beninger RJ (1989b). Dissociating the effects of altered dopaminergic function on performance and learning. *Brain Res Bull* **23**: 365–371.

Berridge KC, Robinson TE (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Rev* **28**: 309–369.

Cardinal RN, Parkinson JA, Hall J, Everitt BJ (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal coretx. *Neurosci Biobehav Rev* **26**: 321–352.

Connell PH (1958). *Amphetamine Psychosis*. Chapman and Hall: London.

Courvoisier S (1956). Pharmacodynamic basis for the use of chlorpromazine in psychiatry. *Q Rev Psychiatry Neurol* **17**: 25–37.

Dayan P, Balleine BW (2002). Reward, motivation and reinforcement learning. *Neuron* **36**: 285–298.

Feldon J, Shofel A, Weiner I (1991). Latent inhibition is unaffected by direct dopamine agonists. *Pharmacol Biochem Behav* **38**: 309–314.

Fibiger HC, Zis AP, Phillips AG (1975). Haloperidol-induced disruption of conditioned avoidance responding: attenuation by prior training or by anticholinergic drugs. *Eur J Pharmacol* **30**: 309–314.

Garris PA, Christensen JR, Rebec GV, Wightman RM (1997). Real-time measurement of electrically evoked extracellular dopamine in the striatum of freely moving rats. *J Neurochem* **68**: 152–161.

Garris PA, Kilpatrick M, Bunin MA, Michael D, Walker QD, Wightman RM (1999). Dissociation of dopamine release in the nucleus accumbens from striatal self-stimulation. *Nature* **398**: 67–69.

Grace AA (1995). The tonic/phasic model of dopamine system regulation: its relevance for understanding how stimulant abuse can alter basal ganglia function. *Drug Alcohol Depend* **37**: 111–129.

Grace AA (1991). Phasic *vs* tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* **41**: 1–24.

Grace AA, Bunney BS, Moore H, Todd CL (1997). Dopamine-cell depolarization block as a model for the therapeutic actions of antipsychotic drugs. *Trends Neurosci* **20**: 31–37.

Gray NS, Pickering AD, Hemsley DR, Dawling S, Gray JA (1992). Abolition of latent inhibition by a single 5 mg dose of d-amphetamine in man. *Psychopharmacology (Berlin)* **107**: 425–430.

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

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

Janssen PAJ, Niemegeers CJE, Schellekens KHL (1965). Is it possible to predict the clinical effects of neuroleptic drugs (major tranquilizers) from animal data? *Arzneimittelforschung* **15**: 104–117.

Joseph MH, Datla K, Young AMJ (2003). The interpretation of the measurement of nucleus accumbens dopamine by vivo dialysis: the kick, the craving or the cognition. *Neurosci Biobehav Rev* **27**: 527–541.

Joseph MH, Peters SL, Moran PM, Grigorvan GA, Young AM, Gray JA (2000). Modulation of latent inhibition in the rat by altered dopamine transmission in the nucleus accumbens at the time of conditioning. *Neuroscience* **101**: 921–930.

Kapur S (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology and pharmacology in Schizophrenia. *Am J Psychiatry* **160**: 13–23.

Kapur S, Mamo D (2003). Half a century of antipsychotics and still a central role for dopamine D2 receptors. *Prog Neuropsychopharmacol Biol Psychiatry* **27**: 1081–1090.

Kapur S, VanderSpek SC, Brownlee BA, Nobrega JN (2003). Antipsychotic dosing in preclinical models is often unrepresentative of the clinical condition: a suggested solution based on *in vivo* occupancy. *J Pharmacol Exp Ther* **305**: 625–631.

Killcross AS, Dickinson A, Robbins TW (1994). Amphetamine-induced disruptions of latent inhibition are reinforcer mediated: implications for animal models of schizophrenic attentional dysfunction. *Psychopharmacology* **115**: 185–195.

Li M, Parkes J, Fletcher PJ, Kapur S (2004). Evaluation of the motor initiation hypothesis of APD-induced conditioned avoidance decreases. *Pharmacol Biochem Behav* **78**: 811–819.

Lubow RE (1989). *Latent Inhibition and Conditioned Attention Theory*. Cambridge University Press: Cambridge.

Lubow RE (1997). Latent inhibition as a measure of learned inattention: some problems and solutions. *Behav Brain Res* **88**: 75–83.

Lubow RE (2005). Construct validity of the animal latent inhibition model of selective attention deficits in schizophrenia. *Schizophr Bull* **31**: 139–153.

Lubow RE, Moore AU (1959). Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. *J Comp Physiol Psychol* **52**: 415–419.

Mackintosh NJ (1974). *The Psychology of Animal Learning*. Academic Press: New York.

Mackintosh NJ (1975). A theory of attention: variations in the associability of stimuli with reinforcement. *Psychol Rev* **82**: 276–298.

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

Mirenowicz J, Schultz W (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* **379**: 449–451.
Montague PR, Dayan P, Person C, Sejnowski TJ (1995). Bee foraging in uncertain environments using predictive Hebbian learning. *Nature* **377**: 725–728.
Montague PR, Dayan P, Sejnowski TJ (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* **16**: 1936–1947.
Moser PC, Hitchcock JM, Lister S, Moran PM (2000). The pharmacology of latent inhibition as an animal model of schizophrenia. *Brain Res Rev* **33**: 275–307.
Phillips PEM, Stuber GD, Helen MLAV, Wightman RM, Carelli RM (2003). Subsecond dopamine release promotes cocaine seeking. *Nature* **422**: 614–618.
Rebec GV, Christensen JR, Guerra C, Bardo MT (1997). Regional and temporal differences in real-time dopamine efflux in the nucleus accumbens during free-chioce novelty. *Brain Res* **776**: 61–67.
Redgrave P, Prescott TJ, Gurney K (1999). Is the short-latency dopamine response too short to signal reward error? *Trends in Neurosci* **22**: 146–151.
Redish DA (2004). Addiction as a computational process gone awry. *Science* **306**: 1944–1947.
Robbins TW, Everitt BJ (1996). Neurobehavioural mechanisms of reward and motivation. *Curr Opin Neurobiol* **6**: 228–236.
Robinson DL, Venton BJ, Heien MLAV, Wightman RM (2003). Detecting subsecond dopamine release with fast-scan cyclic voltammetry *in vivo*. *Clin Chem* **49**: 1763–1773.
Roitman MF, Stuber GD, Phillips PEM, Wightman RM, Carelli RM (2004). Dopamine operates as a subsecond modulator of food seeking. *J Neurosci* **24**: 1265–1271.
Romo R, Schultz W (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *J Neurophysiol* **63**: 592–606.
Salamone JD, Cousins MS, Bucher S (1994). Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. *Behav Brain Res* **65**: 221–229.
Schultz W (1998). Predictive reward signal of dopamine neurons. *J Neurophysiol* **80**: 1–27.
Schultz W, Apicella P, Ljungberg T (1992). Response of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J Neurosci* **13**: 900–913.
Schultz W, Dayan P, Montague PR (1997). A neural substrate of prediction and reward. *Science* **275**: 1593–1599.
Seamans JK, Yang CR (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol* **74**: 1–57.
Segal DS, Kuczenski R (1992). *In vivo* microdialysis reveals a diminished amphetamine-induced DA response corresponding to behavioral sensitization produced by repeated amphetamine pretreatment. *Brain Res* **571**: 330–337.
Solomon PR, Crider A, Winkelman JW, Turi A, Kamer RM, Kaplan LJ (1981). Disrupted latent inhibition in the rat with chronic amphetamine or haloperidol-induced supersensitivity: relationship to schizophrenic attention disorder. *Biol Psychiatry* **16**: 519–537.
Tesauro GJ (1994). TD-Gammon, a self-teaching backgammon program, achieves master-level play. *Neural Computation* **6**: 215–219.
Wadenberg MG, Soliman A, Vanderspek SC, Kapur S (2001). Dopamine D2 receptor occupancy is a common mechanism underlying animal models of antipsychotics and their clinical effects. *Neuropsychopharmacology* **25**: 633–641.
Weiner I (1990). Neural substrates of latent inhibition: the switching model. *Psychol Bull* **108**: 442–461.
Weiner I (2003). The 'two-headed' latent inhibition model of schizophrenia: modeling positive and negative symptoms and their treatment. *Psychopharmacology (Berlin)* **169**: 257–297.
Weiner I, Feldon J, Katz Y (1987). Facilitation of the expression but not the acquisition of latent inhibition by haloperidol in rats. *Pharmacol, Biochem Behav* **26**: 241–246.
Weiner I, Lubow RE, Feldon J (1981). Chronic amphetamine and latent inhibition. *Behav Brain Res* **2**: 285–286.
Weiner I, Lubow RE, Feldon J (1988). Disruption of latent inhibition by acute administration of low doses of amphetamine. *Pharmacol Biochem Behav* **30**: 871–878.
Wise RA (1982). Neuroleptics and operant behavior: the anhedonia hypothesis. *Behav Brain Sci* **5**: 39–87.

APPENDIX

SUPPLEMENTARY MATERIAL

All procedures were approved by the animal care committee at the Centre for Addiction and Mental Health, and adhered to the guidelines established by the Canadian Council on Animal Care.

TD Simulation of Conditioned Avoidance

In TD, the Value of each state is intended as an estimate of all future reward following that state:

$$V(t) = r(t) + \gamma^1 r(t+1) + \gamma^2 r(t+2)... \quad (6)$$

where $r(t)$ is the reward associated with the state of time t, and $r(t+1)$ is the reward associated with the following state etc. γ (between 0 and 1) is the *discount factor* that controls the degree to which future rewards are discounted over immediate rewards (eg a dollar tomorrow is worth slightly less than a dollar today). $\gamma > 0.9$ is commonly used, and we arbitrarily select $\gamma = 0.93$. As with the number of interval states, we select a value that empirically matches the quantitative properties of the observed data, without affecting the qualitative performance of the model.

The Value of each state must be estimated by repeated exposure to that state over multiple trials. To avoid the need to wait indefinitely for all future rewards before updating $V(t)$, TD makes use of the recursive definition:

$$V(t) \approx r(t) + \gamma V(t+1) \quad (7)$$

appealing to the ability of $V(t+1)$ to itself estimate $r(t+1) + r(t+2) +$ etc. The 'temporal difference' between the left and right hand sides of Equation (7) yields an error term, which is then used to update $V(t)$ towards $r(t) + V(t+1)$ (see Equation (1)). The assumption is that $r(t) + V(t+1)$ is a more accurate estimate of the true future reward because it incorporates real reward from the environment (ie $r(t)$). These theoretical foundations of TD were formulated long before the model's application to the dopamine system.

In the simulation of conditioned avoidance, four internal states were used, along with a learning rate, $\alpha = 0.5$, and discount factor $\gamma = 0.93$. The drug free sessions were simulated by setting the 'dopamine' parameter, $\theta = 0$. Low, medium, and high doses of APD were simulated by setting $\theta = -0.2$, $\theta = -03$, $\theta = -0.4$, respectively. The seven sessions of Figure 2 were modeled as $7 \times 30 = 210$ consecutive simulated trials, with only θ being varied from session to session as appropriate.

We note that TD is usually applied to appetitive tasks, with positive prediction error being associated with 'better than expected' and negative prediction error with 'worse than expected'. This presents a problem for modeling aversive paradigms. Although dopamine is clearly implicated in such paradigms, APD-induced dopamine blockade does not appear to render the CS 'worse' or more aversive, but apparently less aversive. Our solution involves treating the shock in the same way as rewards have been treated in the past—by setting r(US) to a positive value. The tacit assumption being made is that the dopamine system mediates the learning of all salient events, whether appetitive or aversive. Electrophysiological evidence of dopamine neuron selectivity for appetitive events (Mirenowicz and Schultz, 1996) undermines this assumption, although microdialysis and voltammetry studies suggest that dopamine is also released in response to aversive events (Joseph *et al*, 2003), and indeed a more general relationship between dopamine and salience has been posited (Horvitz, 2000). The resolution of this debate will have important implications for how aversive paradigms such as CA are modeled using TD in the future.

Latent Inhibition Animal Experiment

The following describes experiment 2 in (Weiner *et al*, 1988). Seventy-two male Wistar rats served as subjects. They were randomly assigned to one of the eight experimental conditions in a $2 \times 2 \times 2$ design, consisting of stimulus pre-exposure/no pre-exposure, drug/no drug during pre-exposure, and drug/no drug during conditioning. The conditioned stimulus was a 5-s, 2.8 kHz tone and the US was a shock (1.0 mA) supplied to the grid floor. Each animal in the pre-exposed group (PE) was placed in the shuttle box and received 50, 5-s tone presentations without consequence. The non pre-exposed (NPE) animals were confined to the shuttle box for an identical period of time, but were not presented with the tone. Twenty-four hr after pre-exposure, each animal was placed in the shuttle box and received 60 trials of avoidance training. Each avoidance trial started with a 5-s tone followed by a 30-s shock, the tone remaining on with the shock. The total number of avoidance responses (shuttle crossings) was recorded. The appropriate drug, either 1.5 mg/kg dl-amphetamine sulphate dissolved in 1 ml of isotonic saline or an equivalent volume of saline, was administered intraperitoneally 15 min prior to the start of each stage (pre-exposure and conditioning). Subsequent reviews (Moser *et al*, 2000; Weiner, 1990) suggest that the conditioning phase rather than the pre-exposure phase is the important stage for pharmacological manipulation of LI, and so we focus on the groups that received saline during pre-exposure.

Supplementary Information accompanies the paper on the Neuropsychopharmacology website (http://www.nature.com/npp).