

Dopamine, Learning, and Impulsivity: A Biological Account of Attention- Deficit/Hyperactivity Disorder

Jonathan Williams, MBBS, MSc, MRCPsych,¹ and Peter Dayan, Ph.D.²

ABSTRACT

Background: Attention-deficit/hyperactivity disorder (ADHD) affects up to 10% of school-age children. The impulsivity which is seen as its core feature persists over years, yet experimental measures of impulsivity can be altered in a single session. In this study, we tested the theory that both the persistence and the variability of impulsivity could be the result of abnormalities in learning mechanisms and environment.

Method: We extended an existing model of the role of dopamine in operant conditioning to address the delayed response time task, which is one of the standard tests for impulsivity in ADHD. In this task, subjects choose between immediate responding for a small reinforcer and later responding for a larger one. We studied the influence on impulsivity of four key parameters of the model: The learning rate, discount factor, brittleness, and action bias.

Results: The behavior of the model is broadly comparable with electrophysiological (monkey) and behavioral (ADHD and normal) data. Variations in any of the parameters can cause impulsivity. All parameters except the discount factor show inverted U-shaped curves for their effects on impulsivity, suggesting, for example, how either hyper- or hypofunctioning of dopamine can cause impulsivity. The model suggests how decision making can be affected by environmental unpredictability, and thus offers an account of one aspect of the natural history of ADHD.

Conclusions: Some types of ADHD may be caused by specific deficits in reinforcement learning and in the use of learned lessons. Environmental factors can interact with these deficits to delay maturation.

INTRODUCTION

ATTENTION-DEFICIT/HYPERACTIVITY DISORDER (ADHD) is a developmental disorder defined as involving difficulties with sustained

attention, hyperactivity, and impulsivity. It affects 5%–10% of school-age children, in severe cases putting their social and psychological development at risk (Scahill and Schwab-Stone 2000; Taylor 1994).

¹Department of Child and Adolescent Psychiatry, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, UK.

²Gatsby Computational Neuroscience Unit, University College London, UK.
Financial support from University College London Department of Psychiatry and Behavioural Sciences and London Postgraduate Medical Deanery, London, United Kingdom.

Financial support from the Gatsby Charitable Foundation, London, United Kingdom.

Neuropsychological deficits have been described in ADHD, particularly in tests of selective attention and frontal function (Doyle et al. 2000; Grodzinsky and Barkley 1999; Lockwood et al. 2001). Reduced activation of various frontal areas has been described in ADHD children during Stroop, stop, and motor-timing tasks (Bush et al. 1999; Rubia et al. 1999). Abnormalities of event-related potentials in a continuous performance task, of visuomotor perception, and of verbal memory and learning have also been described (Oie and Rund 1999; Raggio 1999; Sunohara et al. 1999). Children with ADHD also suffer from academic impairments (Barkley et al. 1991; Faraone et al. 1993). Some impairments persist from preschool to college age (DuPaul et al. 2001; Heiligenstein et al. 1999). These characteristics may be more marked in certain subgroups (Swanson et al. 2000a).

Even though several large twin studies have established the inheritability of ADHD to be approximately 0.80 (Swanson et al. 2000b; Thapar et al. 1999), it is also subject to important environmental influences, both within and between experimental episodes. For example, the learning of ADHD children is particularly susceptible to disruption by noncontingent and partial reinforcement schedules (Douglas and Parry 1983). Slusarek et al. (2001) found that ADHD children's performance in a stop-signal task was deficient under conditions of low incentive but normal with higher incentive. Castellanos et al. (2000) have found that the oculomotor go-no go task, in which ADHD children make twice as many commission errors, and three times as many intrusion errors, as normals, is subject to a "practice effect" sufficient to prevent the use of repeat testing in experiments. Similarly, Chee et al. (1989) felt their data indicated that practice was an important variable affecting performance of ADHD children in a continuous performance task. There is, however, no evidence that such experimental changes generalize to day-to-day life.

ADHD symptoms, though sometimes lasting into late adolescence or adulthood, do appear to reduce with age in many cases (Mannuzza and Klein 2000; Pineda et al. 1999). Hence, both continuity and its apparent opposite, context specificity, are seen in ADHD, in both everyday behavior and the laboratory. This paper

proposes that both aspects are predictable, and result from: (1) idiosyncratic, genetically determined learning mechanisms, and, to a lesser extent, (2) idiosyncratic environments. This conforms with the framework of nature-nurture interactions, which are being increasingly recognized as central to child psychiatry (Rutter and Plomin 1997; Rutter and Rutter 1993).

Impulsivity in ADHD

Many researchers have suggested that the central deficit in ADHD is impulsivity (Barkley 1997; Tannock 1998). Impulsivity generally means acting with inadequate thought, or without adequate consideration of reward or punishment. Varieties of impulsivity can be defined, often based on hypothesized underlying mechanisms (such as deficits in response inhibition or switching or timing). Experimental tests used to access these include delayed reward, go-no go, stop tasks, response time in the uncertain visual discrimination test, and delayed response time tasks (DRTT) (Evenden 1999; Rubia et al. 2001).

In the DRTT, subjects are offered two different-sized reinforcers (of various different sorts): A small one if they act immediately, or a larger one if they act later. Though this willingness to wait has not yet been completely quantified, it is clear that children with hyperactivity are more likely to respond to get the small immediate reinforcer than are control subjects (Rapport et al. 1986). In this study, we modeled impulsivity in the DRTT, exploring, quantitatively, Taylor's (1994) view that "it cannot be assumed from the cognitive studies so far that we are dealing with a deficit of inhibitory control rather than an alteration in the ways that decisions about inhibition are made." Indeed, as in the model, such decisions may be based on many factors other than the size of the reinforcers (Sonuga-Barke et al. 1998).

Computational modeling of neuromodulation in ADHD

Computational modeling is widely used in psychology and neuroscience as a tool for specifying and testing information-processing accounts of neural function and behavioral data,

and for integrating and accounting for large and disparate bodies of experimental data. O'Reilly and Munakata (2000; see also Dayan and Abbott 2001) provide an excellent overview of this approach.

This paper is primarily concerned with the behavioral effects of variations in constraints (or parameters) controlling learning. The mapping of these parameters to brain structures (as in Fig. 1A) is considered mainly where the experimental data are strongest, namely with the dopaminergic neuromodulatory system originating in the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc). The importance of other neurotransmitters and neuromodulators is considered later.

The model presented in this paper is a simple application to the DRTT of a form of learning called temporal-difference learning, which is standard in a computational field called reinforcement learning (Sutton and Barto 1998). The link from this learning method to dopamine function was originally made by Montague et al. 1996; see also Friston et al. 1994) to account for data on the activity of dopamine cells in the VTA and SNc of macaque monkeys during the learning of an operant conditioning task (Schultz 1998). The idea is that the monkeys are constantly learning to predict future reinforcement within a trial, and that the phasic (not tonic) activity of dopaminergic cells signals mismatches in these predictions. This "error signal" is used directly to control the learning of predictions, and the predictions are used, in turn, to control which actions are selected. The model accounts well for a wide variety of data on the dopamine system in learning (Schultz et al. 1997), for which there is also accumulating evidence in humans (Fried et al. 2001). It has also been used to model human (fMRI) data in reward learning (McClure et al. 2003; O'Doherty et al. 2003, 2004).

MODEL AND METHODS

Our model is based on a computational account of the involvement of dopamine in Pavlovian and operant conditioning. In this account, dopaminergic activity reflects ongoing errors in the subjects' predictions of future reinforcement. These errors are used as signals

for learning appropriate predictions and also, in the operant case, for learning actions that maximize the delivery of rewards.

In the version of the DRTT that we modeled, a buzzer sounded at a fixed time in a trial, and then the subject had multiple opportunities to press a lever. If the lever was pressed shortly after the buzzer, then a small reinforcement was given; if the subject waited for longer, a larger reinforcement was provided. The lever could only be pressed once in each trial, and only after the buzzer. In the simulation, time after the buzzer was discretized into steps (of a length of a few seconds), the small reward (magnitude $r = 1$) was given if the (model) lever was pressed in the second timestep after the buzzer, and the larger reward (magnitude $R = 4$) was provided if the computer chose to wait until the fifth timestep before the lever was pressed.

Figure 1A shows the model. The different timesteps following the buzzer count as separate, distinct states, each represented by a unique pattern of neural activity in the cortex (likely the prefrontal cortex). The basal ganglia (together with affective processing structures, such as the amygdala) learn to associate these states, together with potential actions, with predictions of future reinforcements. The "action" (such as pressing a lever or waiting) which appears likely to produce the largest future reinforcement is, in general, selected by the basal ganglia. The "predicted reinforcement" is used together with information about actual reinforcement to create a "prediction error" signal. This, in turn, is used to alter the predictions that are made when, in future, the same situation or state is encountered again. The prediction error signal models activity of the dopaminergic cells in the VTA and SNc.

Prediction and action learning follow the tenets of temporal-difference learning, which is described in detail in the Appendix. Briefly, note first that learning to get the larger, more delayed reinforcer is quite tricky. For instance, consider the case that a subject, or the model, gets a large reinforcer by randomly happening to press the lever for the first time within a trial at timestep 5. To repeat this feat, not only does the subject need to remember the benefit of pressing the lever at this timestep, but the subject also must remember or figure out *not* to press the lever at timestep 4 (and then time-

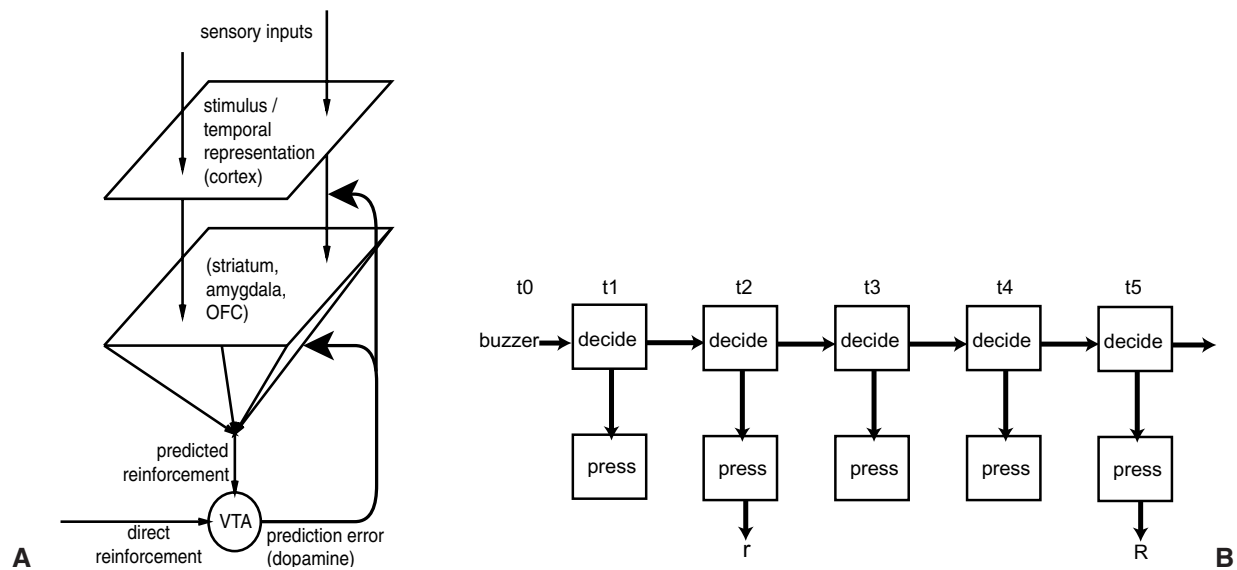


FIG. 1. Overview and operation of the model. (A) Structure of the model. VTA: ventral tegmental area and substantia nigra pars compacta. **(B)** Decisions that must be made by the model in each DRTT trial. At each timestep (t1–t5), the model must select either waiting or pressing the lever. DRTT, delayed response time task; OFC, orbitofrontal cortex.

step 3 and 2 and 1) so that the larger reinforcement is still available. In temporal-difference learning, knowledge about future events usually propagates backward step-by-step over multiple learning trials—in this case, to the time of the buzzer.

To clarify the operation of the model, Figure 1B shows the 10 numbers which the model adjusts to form its representation of the DRTT environment. These numbers comprise a pair of numbers for each timestep, giving the total future reinforcement expected for each of the two choices, namely pressing the lever or waiting. These numbers are represented in what is known as a lookup table, which is the very simplest form of neural network.

Before any trials have been performed, there is no expectation of any reinforcement. The values of the 10 predictions are, therefore, all zero, as follows (corresponding to the structure shown in Fig. 1B):

	t1	t2	t3	t4	t5
Wait:	0	0	0	0	0
Press:	0	0	0	0	0

The following table shows the situation after a hypothetical learning episode in which the model has learned to predict, from the start of any trial, the availability of small reinforcement

r, which can be achieved by deciding at t1 to wait, and at t2 to press. Such a steady state, reflecting ignorance of the availability of large reinforcement R, can be achieved in several ways that we explore later:

	t1	t2	t3	t4	t5
Wait:	1	0	0	0	0
Press:	0	1	0	0	0

The final table shows the situation after a learning episode in which the model has explored his or her options thoroughly and is aware of both r and R; the only real choice facing the model is at t2, when he or she must select between the small reinforcement immediately and the larger one later. The “Press” line reads 0-1-0-0-4 because these are the rewards for pressing at each timestep. Pressing at t1, t3, or t4 is not rewarded at all, and the model correctly learns this.

	t1	t2	t3	t4	t5
Wait:	4	4	4	4	0
Press:	0	1	0	0	4

Figure 2A shows the model learning to perform the task over the course of 500 trials. The short horizontal lines show when, in each trial, the lever was pressed. In the first few trials, the lever is pressed at the time of the buzzer, lead-

ing to no reinforcement. Quite rapidly, the model learns to get the small reinforcement by waiting one timestep after the buzzer, but only after some 250 trials does the model learn to wait for the large, later, reinforcement. Figure 2B shows the prediction error signal that controls learning. Unexpected reinforcers, or “pleasant

surprises,” increase this signal, shown as upgoing peaks in the figure. The color of the peaks is an artefact of the plotting program, but conveniently separates white foreground peaks ($t > 0$) from the black slower-changing values at $t = 0$.

The white peaks form two curved “mountain ranges,” one corresponding to the small rein-

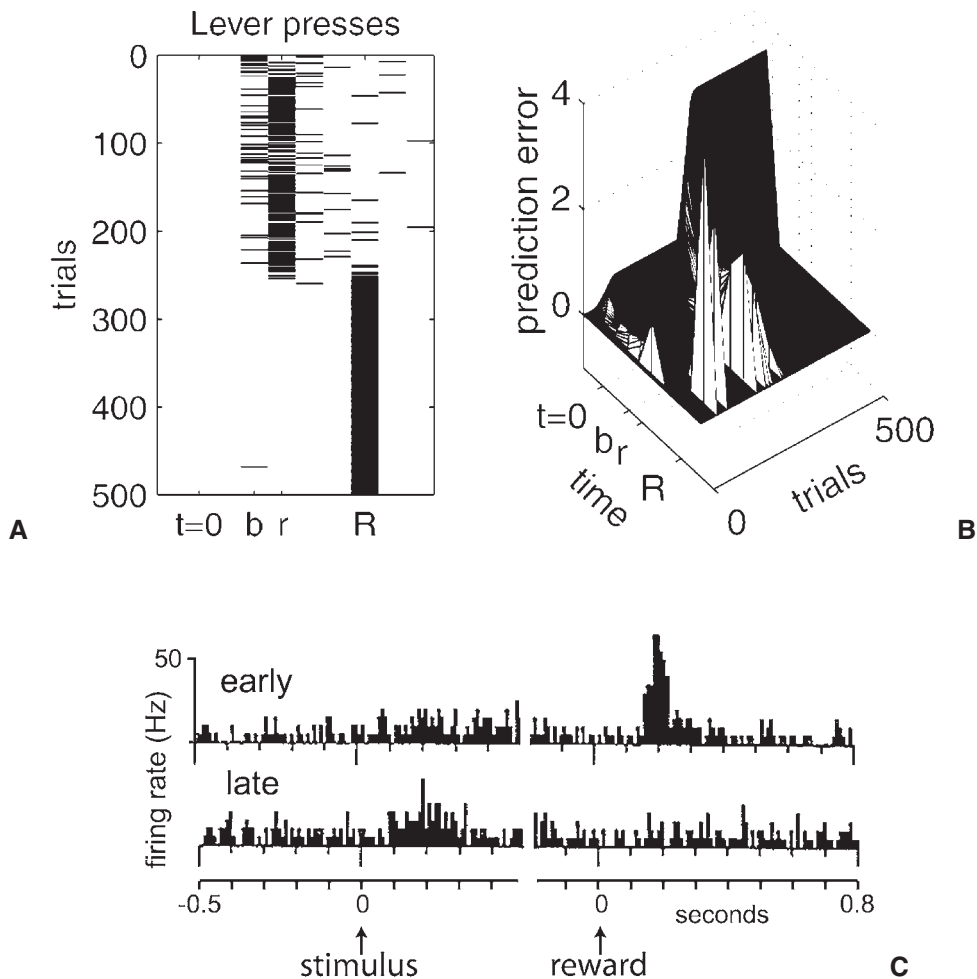


FIG. 2. Comparison with electrophysiological recording in monkey. (A) Choices made by model during normal learning of the task over 500 trials. The horizontal axis starts at $t = 0$ in each trial; b is the time of the buzzer, indicating the start of the period during which the lever can be pressed; r indicates the time when a small reinforcement can be obtained by pressing the lever; R indicates the time when a large reinforcement can be obtained. (B) Prediction error during the same 500 trials. The trial number is shown on one axis, and the time within each trial on another, using notation as in (A). Note that positive prediction errors (shown on the vertical axis) occur earlier within later trials, resulting from lessons accumulated gradually over many trials (see text for explanation). (C) Macaque VTA dopamine cell activity at an early stage (upper plot: corresponds with early trials in (B)) and late stage (lower plot: corresponds with late trials in (B)) of learning to perform an operant conditioning task. Standard values of parameters for A and B and subsequent figures (explained later): brittleness = 2; action bias = 0; discount factor = 0.95; learning rate = 0.3. See text for further explanation. (C) adapted from Mirenovicz and Schultz (1994)). VTA, ventral tegmental area.

forcements, which are increasingly well predicted (so the peaks become smaller and move toward $t = 0$) as the first 20 trials pass; and the other corresponding to the large reinforcements (which are much less frequent in early trials and which also move toward $t = 0$). Figure 2A shows that, within approximately 20 trials, the model is reliably achieving the small reinforcement r . At the same trial in Figure 2B, the small reinforcement r is perfectly predicted from the beginning of the trial ($t = 0$), so the signal, indicating "pleasant surprise," shows a brief positive deflection at the beginning of each trial. After approximately 250 trials, though, the large reinforcement R is reliably achieved (Fig. 2A), so the large reinforcement can be perfectly predicted from the beginning of the trial, and so the error signal at $t = 0$ reliably rises to 4 (Fig. 2B). This signal formally arises from the assumption that the start of each trial is completely unexpected; its existence is crucial for the way that temporal-difference learning models an exactly equivalent signal seen in recordings of dopamine-cell activity (Schultz et al. 1997) and also temporal phenomena in classical conditioning, such as secondary conditioning (Dickinson 1980).

In cases discussed below, the model can get trapped pressing the lever at timestep t_2 and receiving the small reward rather than waiting until t_5 and getting the large reward. Pressing the lever too early is an operational definition of impulsivity. Thus, in studying what controls this outcome for the model, we studied the conceptual provenance of impulsivity.

Parameters of the model

Whether or not the model ever learns to wait for the larger reinforcement, and how the model does so, are controlled by four key parameters. In this context, parameters are numerical measures of long-term aspects of the model's (and thus, putatively, the child's) behavior. These parameters are generally unchanging during a single experimental episode and may be genetically controlled. Although the parameters interact to determine the overall behavior of the model, we studied primarily the simple case of their effects in isolation.

The first parameter of the model is the brittleness, defined as "the extent to which behavior is based on learned lessons" (Holland 1986; Servan-Schreiber et al. 1990 discuss the same idea, using the term "signal-to-noise ratio"). This determines how differences in the predictions of the future reinforcement translate into differences in the propensity of the model to select those actions. Clearly, actions that are expected to lead to higher rewards should be chosen more frequently. Just how much more frequently is determined by the brittleness parameter. Brittleness is one way of controlling the balance between exploitation of existing knowledge (in the difference in predicted rewards) and exploration to improve and refine the knowledge. Exploration becomes more important in "noisy" and changing environments, in which the past is only an imperfect guide to the future. If the model is set to be very brittle, then early observation of the small reward r will make pressing at t_2 overwhelmingly more likely than waiting. This, in turn, will make it hard for the model ever to discover the large reward R , and will, therefore, impede the developmental progression of the model beyond the type of behavior characterized as impulsive. If the brittleness is set very low, then the model would exhibit a comparative inability to persist with one behavior, even when it had collected adequate information about reinforcement availability. This effect is closely related to Servan-Schreiber et al.'s (1990) analysis of the influence of stimulants on the ability of networks to detect a signal embedded in noise.

The second parameter is action bias, which is a measure of the model's preference of action over inaction. This is greater than zero if there is an innate bias to act, as would arise if the action of pressing is, itself, reinforcing. Conversely, it is negative if inactivity is preferred. A non-zero action bias can force the model to make decisions which are suboptimal from the perspective of harvesting external rewards. Such a preference could, in theory, be innate or learned, or both.

The third parameter is learning rate. In the model, changes in predictions (and, thereby, changes in the probabilities of actions) are proportional to the prediction error. The learning

rate is the constant of proportionality and, therefore, has a multiplicative effect on the speed of change. Note that for consistency with the computational literature, the term is taken to mean the maximal rate at which predictions can change, rather than, as would be more natural from a behaviorist perspective, the rate at which behavior changes. Changing the learning rate has significant effects beyond simply altering the rate of DRTT acquisition. For instance, if learning happens too quickly, then the merits of the first reinforcement found (which is likely to be the smaller one) can be learned so strongly that the model never explores adequately to find the other reinforcement. Conversely, if the learning rate is too slow, then it may take too long (more than 500 trials in our model) to learn about the later reinforcement. In our study, we treated the learning rate as a surrogate for various heritable factors associated with enhanced or suppressed release of dopamine. However, various other factors are also likely to control the learning rate, notably cholinergic and noradrenergic neuromodulation (Holland and Gallagher 1999; Yu and Dayan 2002a,b).

The final parameter is the discount factor (D). The idea in this parameter is that a reinforcement expected in the future is worth less than the same reinforcement delivered now. We quantify this by multiplying the reinforcement by a number D (between 0 and 1) at each timestep. At timestep t_5 , when it is actually received, the larger reinforcement is worth R . However, at timestep t_4 , the value of this future reinforcement is decreased by a factor of D , reflecting the fact that it won't be available for one timestep. Thus, it will only be worth RD . At timestep t_3 , this same reward (available at t_5) is worth even less, namely RD^2 , reflecting the two timesteps that must be waited. At timestep t_2 , the subject faces a choice between a small, immediate reinforcement, worth r , and a late, large reinforcement, worth RD^3 . Using the terminology of our model, the ADHD children have a smaller D , so future reinforcements are discounted more, and they will choose the small reinforcement more often than control children. Indeed, Sagvolden et al. (1998) have explicitly demonstrated this effect, which they describe as a "shorter and steeper delay gradi-

ent" in ADHD children, and they have proposed that this gradient leads to the development of overactivity, increased behavioral variability, motor impulsiveness, and impaired sustained attention (Sagvolden et al. 2000). Such a gradient has also been correlated with impulsive behaviour in general psychiatric outpatients (Crean et al. 2000) and has been quantified in heroin addicts (a disorder sometimes comorbid with ADHD) as twice that of controls (Kirby et al. 1999).

The kind of discounting we used is called *exponential*, meaning that a reinforcement loses a fixed proportion of its value at each timestep. However, a different kind of curve, the hyperbolic, better describes the discounting found in many psychological studies (Ainslie 1975; but see also Kacelnik 1997; Kirby and Herrnstein 1995; Madden et al. 2003; Monterosso and Ainslie 1999). The key difference is that with hyperbolic discounting, preferences between two reinforcements—one early and one late—can reverse as the time to both increases. However, there is, as yet, no evidence that children show preference reversals in the DRTT, and so, for simplicity, we preserve the exponential discounting that Montague et al. (1996) used in their model.

RESULTS

Comparison with actual electrophysiological recording

Figure 2C shows Schultz's (1998) recordings from monkey dopamine neurons, in an operant-conditioning task somewhat similar to ours. In early trials, the dopamine cells show slightly increased firing at the time of the stimulus, and much greater firing when the reinforcement is delivered. However, in late trials (i.e., after learning), there is no response to the reinforcement, but only to its earliest reliable predictor, the stimulus. The "early" recording in Figure 2C can be compared with trials 1–100 in Figure 2B, and the "late" recording with trial 500 in Figure 2B. Both the experimental data and the simulation show a movement of excitation, from the time of the reinforcer in early trials, toward its

earliest reliable predictor in later trials. The model's stimulus-linked signals (at $t = 0$) are large because very little discounting is used in this simulation.

Comparison with actual delayed reward performance in ADHD

Figure 3 shows the performance of the model in a paradigm similar to the DRTT, but with a single decision rather than the five shown in Figure 1B. These results can be compared with those from real children (Sonuga-Barke et al. 1992). The "trials constraint" part of that study, in which the children had a fixed number of trials, is the closest to ours. The experimenters explained the fixed number of trials thoroughly to the children before they started. Over 20 trials, 18% of delayed large rewards (standard deviation,

20%) were obtained by hyperactive children, compared to 48% (standard deviation, 34%) obtained by controls. Their data match the model's, if one arbitrarily assumes that their thorough explanation had an effect comparable to preliminary trials (we had no way of verbally priming our model, and preliminary trials are often included in such studies, e.g., 30 in Slusarek et al. (2001)). The possibility that task performance had not stabilized by the end of the experiment is supported by: (1) the large standard deviations, and (2) the fact that this experiment had fewer trials than several other published trials, which mentioned major practice effects in other paradigms (Castellanos et al. 2000; Chee et al. 1989; see also Tannock and Schachar 1992).

For the lower learning rate shown in Figure 3, the model's performance in "trials constraint"

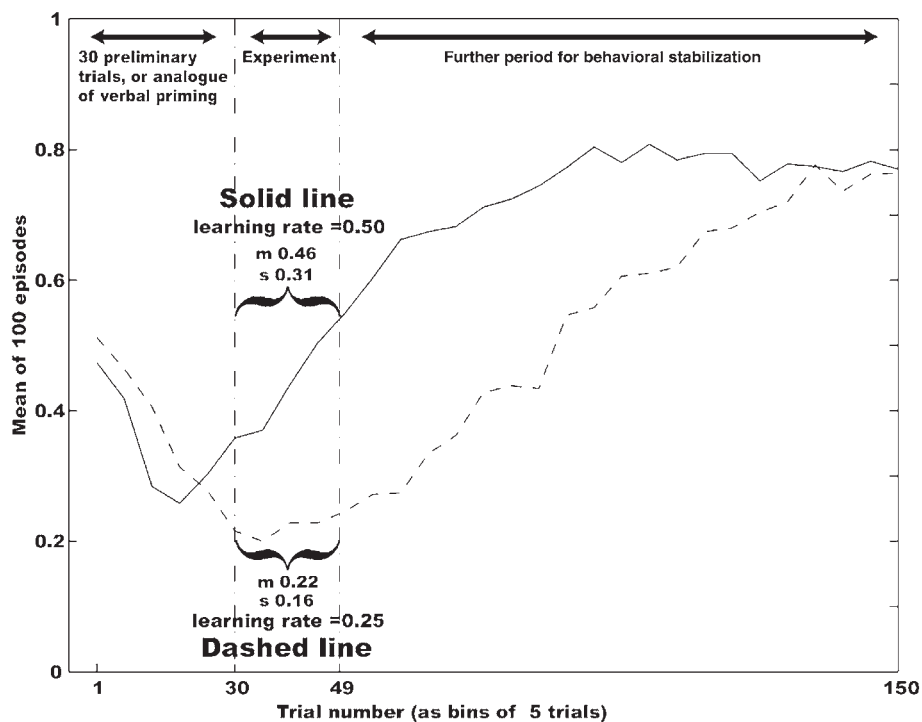


FIG. 3. Effect on model's performance of altering the learning rate. The lines were obtained by averaging 100 simulated learning episodes, each consisting of 150 trials in which the model had a single choice, between a small immediate reinforcement ($r = 1$) and a larger delayed reinforcement ($R = 2$). This paradigm is simpler than that used in the rest of the current paper. The solid line was made using a learning rate of 0.5. The dotted line was made using a lower learning rate (0.25) as one candidate explanation for impulsivity. Apart from the learning rate, parameters are as in Figure 2. (m = mean; s = standard deviation). Simulated results in the figure can be compared with actual results from Sonuga-Barke et al (1992): hyperactive children achieved R on 18% of trials (standard deviation 20%), whereas control children scored 48% (standard deviation 34%). Note that local stability of performance in the experimental period does not reliably indicate achievement of asymptotic performance.

learning is altered, but not its asymptotic performance after prolonged learning. In general, changes in any of the parameters are able to influence both “trials constraint” and asymptotic performance. Brittleness and action bias influence behavior directly, whereas learning rate and discounting act more indirectly by affecting the gradual acquisition of predictions.

Effect on impulsivity of varying learning parameters

Figure 4 shows the basic effect of changing the parameters away from their standard values. Each solid line in Figures A–D was produced by

varying a single parameter. The first three of these are inverted-U-shaped curves, but the fourth (discount factor) is not. Interestingly, U-shaped dose-response curves have previously been predicted to be particularly common in task acquisition (Tannock and Schachar 1992).

We also tested the effect of interactions between the above manipulations and simultaneously increasing the learning rate (dotted lines in Figures A–D). Given the parallel between learning rate and dopamine in our model, it is unsurprising that Figure 4C displays a prediction that prodopaminergic agents will reduce impulsivity caused by preexisting hypodopaminergia (as the dotted treatment line is higher than the solid un-

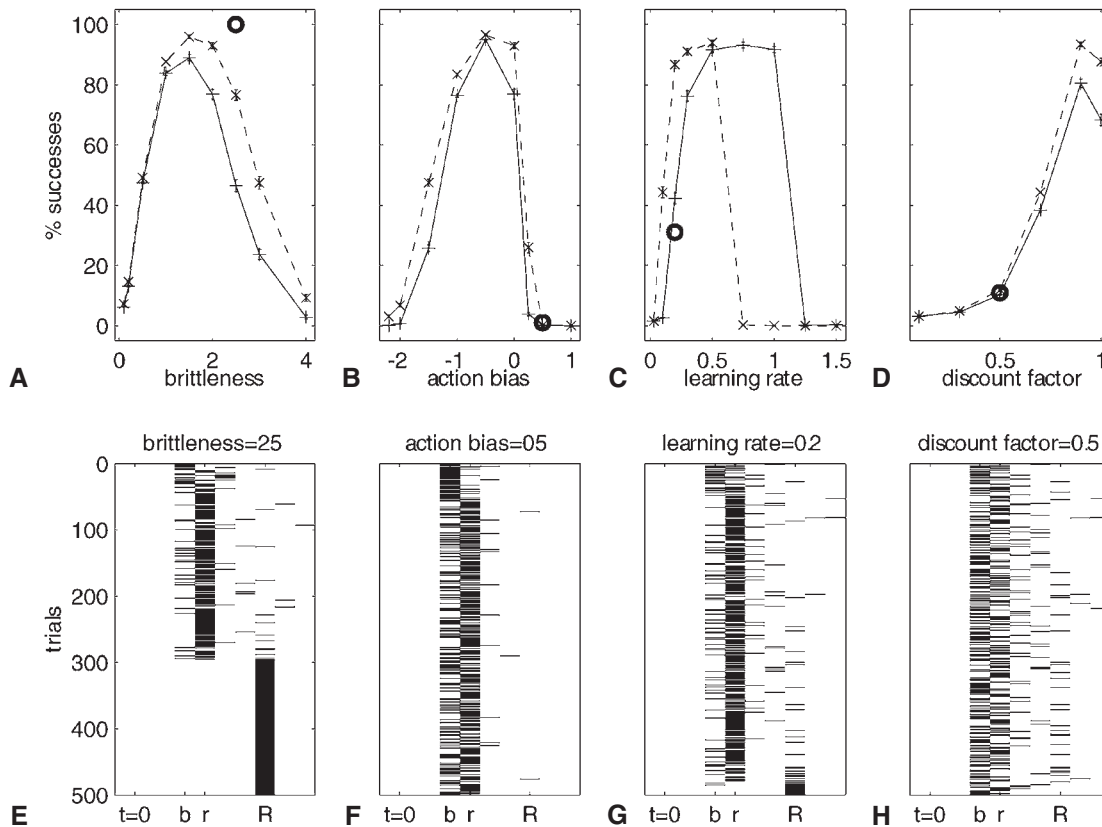


FIG. 4. Factors affecting model's performance on DRTT. The four columns show the model's behavior, using the same parameters as in Figure 2, except that in each column the effect of changes in one parameter are shown. In (A), the effect on performance of varying brittleness between 0 and 4 is shown by the dark line, with "+" showing each datapoint which is the percentage of large reinforcements R achieved in trials 401–500, averaged over 500 episodes. Error bars indicate ± 1 standard error of the mean. The circle indicates the result for one episode randomly chosen from the 4500, in this case with brittleness = 2.5, and shown in full in (E). In (A), as in (B–D), the effect on performance of doubling the learning rate is shown by the dotted lines, with "x" at each datapoint. (B) shows the effect of varying the action bias, (C) the learning rate, and (D) the discounting. Because the x-axis of (C) is the learning rate, comparison of the solid and dotted lines in (C) provides a trivial verification that the program is correctly doubling the learning rate. This can also be interpreted as the effect of dopaminergic genes on performance (solid line), being modified by hypothetical prodopaminergic medication (dotted line). This medication improves performance for subjects starting in the left half of the plot, but has a negative effect on those starting in the right half. DRTT, delayed response time task.

treated line when the learning rate is low). On the other hand, the model predicts that such agents will not reduce the impulsivity caused by hyperdopaminergia, if such a thing exists (with the learning rate over 0.5, the solid line is higher than the dotted one.) Interestingly, benefit from simulated “treatment” is seen for all the factors. This is a possible explanation for the apparent paradox that while ADHD has many causes, stimulants help in most of them.

Figures 4E–G show that particular values of parameters produce distinct “behavioral styles,” including various degrees of repetitiveness. Severe and degenerative cognitive disorders are often associated with repetitive activity (and with reduced learning). A characteristic of this perseveration is that individuals repeat their last utterances. However, the current model does not explicitly address such perseveration, as during any trial, the model retains no information about the preceding trial, distinct from other trials, and so cannot preferentially repeat the actions performed in the previous trial. The repetitiveness seen in the DRTT task is caused by repeated *independent* action selections, all based on a relatively stable knowledge of the task.

Brittleness

Figure 4A shows an inverted U-shaped relationship between brittleness and performance.

Figure 5 shows the same effect but in more detail.

A priori, one might expect that impulsive children had more random behavior (i.e., less brittle decision making). The model shows that this is not necessarily the case. Note that in Figures 5C and D, simulations with moderate brittleness achieve R consistently within 250–350 trials. Also as expected, in the insufficiently brittle simulations of Figures 5A and B, the model errs on the side of exploration, not sufficiently using the information it already has about reinforcement availability. However, contrary to the a priori view, the highly brittle, over-regulated model in Figure 5E demonstrates “impulsivity” as well. This variety of impulsivity is quite different from that seen in Figures 5A and B: After finding the small reinforcement, the model concentrates its effort on that small reinforcement, never exploring enough to find the large reinforcement.

Figure 5F plots the entropy (i.e., the randomness; see the Appendix) of the responding as a function of brittleness. As seen in Figures 5A–E, the entropy decreases with increasing brittleness.

Effect of environment on impulsivity

As we mentioned above, when $RD^3 < r$, it is appropriate to choose the small reinforcement instead of waiting for the large one. Changes

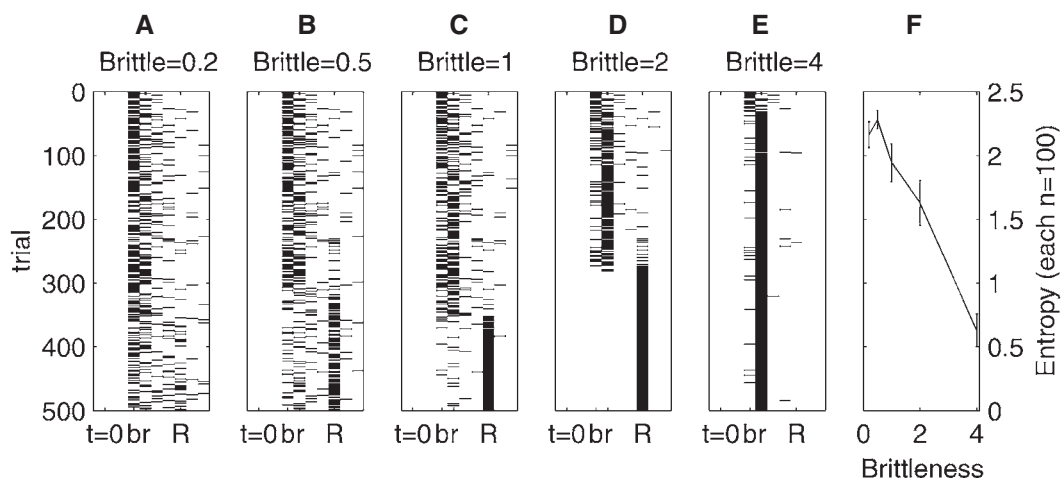


FIG. 5. Brittleness. (A–E) show that increasing brittleness produces first an improvement (B,C) and then a deterioration (D,E) in DRTT performance. (F) shows the means and standard deviations of the entropy, in bits, of the timestep at which the lever is pressed, for values of brittleness as in (A–E), over 100 runs (see the Appendix). DRTT, delayed response time task.

in D affect the model's performance, as shown in Figure 4D, and it is interesting to see how D might be environmentally determined (i.e., subject to learning; see the Appendix for details).

One theoretical interpretation of D is as the probability that the whole trial will unexpectedly terminate at each timestep, thus denying the subject the late reward that it might be awaiting. In Figure 6, we consider the possibility that D might "learn" the predictability of the environment, thus altering the apparent impulsivity over time. Although this effect is in-

teresting, the longitudinal simulation shown in Figure 6 is based on rather crude assumptions, notably that the only memory for the task preserved between learning episodes is actually D . Furthermore, D only changes during the DRTTs, which are modeled as presented every year for 5 years; in reality, we imagine that it will change during the multitude of other tasks that real children are faced with during each year. The graphs demonstrate that as the model grows "older" (i.e., learns to discount less), the model becomes faster at achieving optimal re-

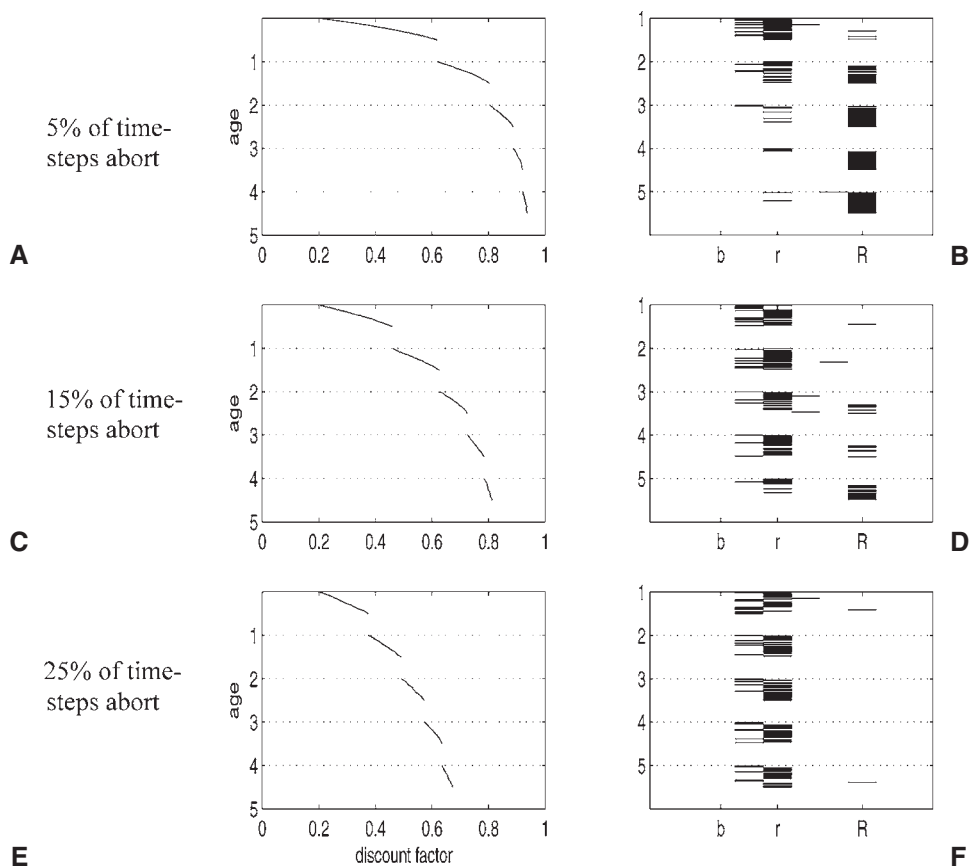


FIG. 6. Model of environmental effects on impulsivity. Each of the three rows in this figure depicts a "virtual child" tested on his first, second, third, fourth, and fifth birthdays, in 500 trials of the DRTT. The three children start identically, with discount factor = 0.2, but are exposed to environments of differing reliability. These environments are modeled by making the task itself somewhat unpredictable, and steadily adjusting the discount factor during trials, to track success in predicting rewards (see the Appendix for details). The first child (Figures A and B) has a substantially predictable environment, in which during only 5% of timesteps is the current trial interrupted. (A) shows the development of the discount factor over trials; (B) shows the resulting performance on the annual test. This child acquires, by age 3, the ability to delay his responding easily in order to achieve a larger reinforcement. Graphs C–F show the consequences of environments in which there is either a 15% or a 25% chance per timestep that each trial will terminate. In either case, the development of delayed responding is abnormal, being either slowed (C,D) or practically absent (E,F). For clarity, only each tenth response is plotted. DRTT, delayed response time task.

sponse delay in the DRTT, but that this is subject to a major effect from the predictability of the environment.

DISCUSSION

Principal findings

We have used a pre-existing learning model to explore the conceptual roots of impulsivity in ADHD. We have validated the model, qualitatively or quantitatively, against genetic, electrophysiological, neuropsychological, pharmacological, and developmental data. The behavior of the model reflects the action of a number of parameters, which we have given neurobiological and psychiatric interpretations. In particular, we have studied the possible role played by dopaminergic mechanisms. The model shows that variation in simple learning and behavioral parameters can all lead to impulsivity in the DRTT. This supports our underlying theory, that ADHD can be caused by abnormalities in learning mechanisms, in the use of learned information, or in information available to be learned. In the model, apparent impulsivity in a DRTT can be caused by overregulated or underregulated behavior, an innate action bias, a hypo- or hyperfunctioning prediction error signal, or by discounting of delayed reinforcements (see Fig. 4).

One interesting and counterintuitive suggestion from the model is that there may be people with too high a learning rate (i.e., too high to allow them to achieve optimal performance on a particular task, as at the right of Figure 4C). Such people would be expected to adapt quickly, and to place inadequate weight on old lessons (Tripp and Alsop 1999). Note that learning that was too rapid for optimal performance on varying or multichoice tasks could *aid* performance on simple tasks, and so might not impair the functioning of such an individual in a simple environment.

An inability to delay responding is often described as a deficit in response inhibition. Response inhibition has been suggested as forming part of an executive function system, in which defects have been documented in

ADHD (Barkley 1997; Bayliss and Roodenrys 2000; Shallice and Burgess 1996), which does not form part of our model. However, noninhibitory deficits may be just as prominent (Shallice et al. 2002), and it has been suggested that deficits in executive function may characterize particular subtypes of ADHD. For instance, Swanson et al. (2000a) found a subgroup comprising 40.6% of their ADHD group (and characterized by a 7-repeat D4 polymorphism), within which, at a mean age of 11 years, psychological testing revealed no deficits on response inhibition, re-engagement, shifting and maintenance of attention, or conflict resolution. (In the other, somewhat larger, subgroup, a proportion of subjects did have such deficits.) If confirmed, this suggests that these executive deficits are not necessary in the development of ADHD. It is possible that the development of executive functions can be either altered by psychosocial effects of impulsivity, or merely delayed so that, by adulthood, the deficit has often disappeared (Walker et al. 2000).

In its current form, our model concentrates specifically on reinforcement. However, dopaminergic systems are also activated by various other factors, notably novelty (Cloninger 1987; Horvitz 2000; Schultz 1998). Novelty can be intrinsically reinforcing—animals will work to deliver novel stimuli that are not primary reinforcers and have not been associated with primary reinforcers (Reed et al. 1996). Novelty may account for the small increase in dopamine firing seen in Figure 2C, in early trials at the time of the stimulus. Such novelty-based responses (sometimes called bonuses in the computational literature (Kakade and Dayan 2002)) play the important role in some abstract reinforcement learning models of controlling the exploration to find optimal actions. Indeed, increased novelty-seeking is found in ADHD (Downey et al. 1996; Young et al. 2000), and associations have been found between a specific allele of the dopamine D4 receptor and novelty-seeking (Malhotra and Goldman 2000).

Our model does not provide a natural explanation for clearly unlearned changes in behavior which rapidly follow stimulant administration, including the “zombie effect” in overdose. We have explored ways of extending our model to

incorporate a “set point” for dopamine levels, which may provide a simple explanation for the behavioral effects of both medication and novelty (Williams and Taylor 2004). For example, either stimulant medication or novel stimuli could produce an immediate increase in dopaminergic transmission; when this exceeded a genetically determined set point, dopamine-seeking behavior would reduce.

By not scaling parameters, we have intentionally avoided the misleading practice of choosing numerical values for parameters in neural networks, giving precise matches with data when the real target is matches with process. Instead, in our case, we are specifically demonstrating how a previously described process (dopaminergic temporal-difference learning) can fit not just one set of data, but a wide range of data (measurements in ADHD). This is, we believe, an important way of using computational models, particularly for disorders such as ADHD, which are on a continuum with normality.

We should emphasize that this paper is intended to present a new approach to studying ADHD rather than being a complete theory of the disorder. Such a theory would need to account for performance differences between International Classification of Diseases (ICD) and Diagnostic and Statistical Manual (DSM) types and subtypes of hyperactivity, as well as explaining why only certain subtypes have associated executive function deficits (Swanson et al. 2000a). Although we have specified four parameters which can control impulsivity, specific values for these, in various subtypes of ADHD, cannot be specified until more detailed experimental data become available. ADHD deficits in a wide range of testing paradigms need to be explained, beyond the two paradigms explored in this paper. The model does not address frustration or anger, because the electrophysiological and computational understanding of these paradigms is currently less advanced than that of dopamine. Sequelae of ADHD may include anger, but our model focuses more on what seem, to us, to be earlier etiological mechanisms.

Furthermore, a complete theory will need to specify the roles of transmitters other than dopamine. Several serotonergic and noradrenergic

genes have been implicated in ADHD (Comings et al. 2000). Serotonin is particularly important in theories of impulsivity, and may have particular relevance to learned and aversive aspects (Carter and Pycocck 1978; Meneses 1998; Popper 1997). Also implicated are noradrenaline, glutamate, and possibly acetylcholine, GABA, and enkephalin (Blum et al. 2000; Carlsson 2001; Comings et al. 2000). It is our hope that the parameters of the model will suggest novel functional roles for these transmitters.

Of course, a weight of evidence implicates dopamine in ADHD, captured in terms of “learning rate” in the model. A variant D4 receptor repeat length and a variant DAT1 transporter allele have been associated with ADHD (Cook, Jr. et al. 1995; LaHoste et al. 1996). Furthermore, siblings with higher numbers of DAT1 high-risk alleles have higher symptom levels (Waldman et al. 1998). Elevated midbrain accumulation of dopamine precursors has been found in adolescents with ADHD (Ernst et al. 1999). In addition, symptoms of hyperactivity and inattention can be produced indirectly by reduction of dopamine transmission, as in phenylketonuria (Diamond 1998; Weglage et al. 1996), or in rats by selective lesions of the VTA (Koob et al. 1989). ADHD can be treated by prodopaminergic stimulants (see below). Importantly, though, evidence exists that both hyper- and hypofunctioning dopamine systems may be associated with hyperactivity (Sagvolden and Sergeant 1998; Spiewoy et al. 2000), as also seen in our model (Fig. 4C).

Changes during development

Based on an interpretation of discounting in which it reflects the chance of a premature termination of a trial, we have suggested how an environment replete with inconsistent reinforcement schedules might lead to the development of impulsive responding. In this case, impulsive responding is actually an optimal solution to an abnormal environment. By the same reasoning, impulsive responding can be an optimal adaptation to a child’s own deficits in predicting his environment.

We interpreted the discount factor D as coming from the possibility, learned by the subjects over the course of many operant tasks, that trials

will end unexpectedly without provision of the delayed reinforcement. That is, children living in a persistently unpredictable environment should learn to discount more steeply than children living in a normal environment, and this will translate into apparent impulsivity. For example, ADHD is significantly associated with large family size (Eddy et al. 1999); unfortunately, we do not know the effect of birth order, which could shed light on the well-replicated absence of significant additive contributions from shared environmental factors (Levy et al. 1997).

ADHD symptoms decrease as a child grows older (Biederman et al. 2000; Scahill and Schwab-Stone 2000), though some of the apparent decrease may result from the use of age-inappropriate diagnostic criteria. Hyperactive-impulsive symptoms probably decline faster than symptoms of inattention (Pineda et al. 1999). "Developmental lag theory" (McLaren 1989; Sagvolden 1996) predicts that "at each age level studied, children with ADHD would perform like younger children without ADHD" (Barkley 1997). Several subclasses of overactive behavior are also found in normal children, most commonly in the first few years of life (De Negri 1995). We presented a learned reduction in discounting (i.e., a learned *increase* in D) as a possible means by which both normal and ADHD children may gradually learn, or mature, out of the hyperactivity of their infancy.

Comparison with other models

The main existing model is that of Ownby and Harwood (1998), who model the sustained attention deficit in ADHD by teaching a three-layer feed-forward network to perform the continuous performance task. They show that the loss of neural connections (for which there is no evidence in ADHD) could lead to degradation of performance. It would be interesting future work to model this task using an extension of our model.

By comparison with ADHD, which has attracted little modeling, there are many computational approaches to phenomena of attention (Parasuraman 1998; Pashler 1998). Although one could consider modeling ADHD by studying modes of failure of these models, none of

them addresses tasks, like the DRTT, that are validated experimentally in ADHD.

ADHD medication

Stimulants such as methylphenidate (MPH) block dopamine transporters (Volkow et al. 1998), and are used to treat ADHD, in which they normalize various neuropsychological measures and, to a limited extent, improve academic performance (Elia et al. 1993; Haenlein and Caul 1987; Sunohara et al. 1999; Tannock et al. 1989). Despite individual etiological factors, each appearing in a minority of cases of ADHD, the great majority of cases are responsive to MPH, in a clear parallel with the results in Figures 4A–D. Comparable effects of MPH are found in some rat strains (Richards et al. 1999).

The possibility that stimulants act by altering the processing of reinforcement in the brain has experimental support from many studies (Cador et al. 1991; Evenden 1999; see McBride et al. 1999; Richards et al. 1999). MPH appears to increase the reinforcement value of rewards in ADHD (Wilkison et al. 1995). Tannock and Schachar (1992) found that effects of MPH reversed between the first and second assessments of a learning task, and suggested that this might reflect different effects on acquisition and performance. Tonic MPH-induced increase in the dopamine signal may achieve its clinical effect by overcoming the elevated reinforcement threshold postulated as a basis for ADHD by Haenlein and Caul (1987).

Clonidine, a second-tier ADHD treatment, stimulates central alpha-2 receptors, thus reducing release of noradrenaline. MPH also influences noradrenaline, beside its primary effect on dopamine. For instance, in ADHD patients who responded to MPH, urinary excretion of a noradrenaline metabolite was reduced by 43%, suggesting a depressive affect on noradrenergic processing (see also Biederman and Spencer 1999; Shekim et al. 1979). Though the role of noradrenaline in cognitive regulation is well supported (Usher et al. 1999), as is its importance in prefrontal cortex, working memory, and attention (Arnsten 2000), we did not assign noradrenaline a place in the current model because, bar this, there is insufficient evidence of its relation to ADHD. In any case, clonidine is

considerably less effective than stimulants (Abo-Zena et al. 2000; Connor et al. 1999).

CONCLUSIONS

Several falsifiable predictions follow directly from the above results. We predict that the performance of a subgroup of ADHD children on measures of impulsivity will normalize with practice to be asymptotically the same as that of well-matched control children. We predict that some ADHD sufferers are particularly prone to the induction of impulsivity by purely behavioral means. If there is a variety of ADHD caused by hyperdopaminergia, the model predicts, unsurprisingly, that MPH will worsen hyperactivity in this. The behavior of the model's extension, described in Figure 6, suggests that the age at which certain individuals' impulsivity reduces will be related to the predictability of their environments (though accurate quantitation of environmental reliability is currently not practical).

In contrast to the usual view that impulsivity affects learning, we believe we have demonstrated the reverse, that learning mechanisms and learned lessons can both affect impulsivity.

Note that the model is able to learn many delayed-reward tasks (see, e.g., Fig. 3). However, we have focussed on the DRTT for its naturalistic provision of: (1) multiple decision points in each trial, and (2) the potentially relevant cognitive hurdle of making the association between events separated in time, namely, the rewards and the buzzer.

ACKNOWLEDGMENTS

We are very grateful to Professor Tim Shallice, Professor Eric Taylor, and Dr. Michael Thomas for their helpful comments on earlier drafts of this manuscript.

REFERENCES

- Abo-Zena RA, Bobek MB, Dweik RA: Hypertensive urgency induced by an interaction of mirtazapine and clonidine. *Pharmacotherapy* 20:476–478, 2000.
- Ainslie G: Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychol Bull* 82:463–496, 1975.
- Arnsten AF: Genetics of childhood disorders: XVIII. ADHD, Part 2: Norepinephrine has a critical modulatory influence on prefrontal cortical function. *J Am Acad Child Adolesc Psychiatry* 39:1201–1203, 2000.
- Barkley RA: Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychol Bull* 121:65–94, 1997.
- Barkley RA, Anastopoulos AD, Guevremont DC, Fletcher KE: Adolescents with ADHD: Patterns of behavioral adjustment, academic functioning, and treatment utilization. *J Am Acad Child Adolesc Psychiatry* 30:752–761, 1991.
- Bayliss DM, Roodenrys S: Executive processing and attention-deficit/hyperactivity disorder: An application of the supervisory attentional system. *Dev Neuropsychol* 17:161–180, 2000.
- Biederman J, Mick E, Faraone SV: Age-dependent decline of symptoms of attention-deficit/hyperactivity disorder: Impact of remission definition and symptom type. *Am J Psychiatry* 157:816–818, 2000.
- Biederman J, Spencer T: Attention-deficit/hyperactivity disorder (ADHD) as a noradrenergic disorder. *Biol Psychiat* 46:1234–1242, 1999.
- Blum K, Braverman ER, Holder JM, Lubar JF, Monastra VJ, Miller D, Lubar JO, Chen TJ, Comings DE: Reward deficiency syndrome: A biogenetic model for the diagnosis and treatment of impulsive, addictive, and compulsive behaviors. *J Psychoactive Drugs* 32(Suppl):i–112, 2000.
- Bush G, Frazier JA, Rauch SL, Seidman LJ, Whalen PJ, Jenike MA, Rosen BR, Biederman J: Anterior cingulate cortex dysfunction in attention-deficit/hyperactivity disorder revealed by fMRI and the Counting Stroop. *Biol Psychiat* 45:1542–1552, 1999.
- Cador M, Taylor JR, Robbins TW: Potentiation of the effects of reward-related stimuli by dopaminergic-dependent mechanisms in the nucleus accumbens. *Psychopharmacology (Berl)* 104:377–385, 1991.
- Carlsson ML: On the role of prefrontal cortex glutamate for the antithetical phenomenology of obsessive compulsive disorder and attention-deficit/hyperactivity disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 25:5–26, 2001.
- Carter CJ, Pycock CJ: Differential effects of central serotonin manipulation on hyperactive and stereotyped behavior. *Life Sci* 23:953–960, 1978.
- Castellanos FX, Marvasti FF, Ducharme JL, Walter JM, Israel ME, Krain A, Pavlovsky C, Hommer DW: Executive function oculomotor tasks in girls with ADHD. *J Am Acad Child Adolesc Psychiatry* 39:644–650, 2000.
- Chee P, Logan G, Schachar R, Lindsay P, Wachsuth R: Effects of event rate and display

- time on sustained attention in hyperactive, normal, and control children. *J Abnorm Child Psychol* 17:371–391, 1989.
- Cloninger CR: A systematic method for clinical description and classification of personality variants. A proposal. *Arch Gen Psychiatry* 44:573–588, 1987.
- Comings DE, Gade-Andavolu R, Gonzalez N, Wu S, Muhleman D, Blake H, Dietz G, Saucier G, MacMurray JP: Comparison of the role of dopamine, serotonin, and noradrenaline genes in ADHD, ODD, and conduct disorder: Multivariate regression analysis of 20 genes. *Clin Genet* 57:178–196, 2000.
- Cook EH, Jr, Stein MA, Krasowski MD, Cox NJ, Olkon DM, Kieffer JE, Leventhal BL: Association of attention-deficit disorder and the dopamine transporter gene. *Am J Hum Genet* 56:993–998, 1995.
- Crean JP, de Wit H, Richards JB: Reward discounting as a measure of impulsive behavior in a psychiatric outpatient population. *Exp Clin Psychopharmacol* 8:155–162, 2000.
- Dayan P, Abbott L: *Theoretical Neuroscience*. Cambridge (Massachusetts), MIT Press, 2001.
- De Negri M: Hyperkinetic behavior, attention-deficit disorder, conduct disorder, and instabile psychomotrice: Identity, analogies, and misunderstandings. Commentary to Gordons' paper (*Brain Dev* 15:169–172, 1994), *Brain Dev* 17:146–147, 1995.
- Diamond A: Evidence for the importance of dopamine for prefrontal cortex functions early in life. In: *The Prefrontal Cortex: Executive and Cognitive Functions*. Edited by Roberts AC, Robbins TW. New York, Oxford University Press, 1998, pp 144–164.
- Douglas VI, Parry PA: Effects of reward on delayed reaction time task performance of hyperactive children. *J Abnorm Child Psychol* 11:313–326, 1983.
- Downey KK, Pomerleau CS, Pomerleau OF: Personality differences related to smoking and adult attention-deficit/hyperactivity disorder. *J Subst Abuse* 8:129–135, 1996.
- Doyle AE, Biederman J, Seidman LJ, Weber W, Faraone SV: Diagnostic efficiency of neuropsychological test scores for discriminating boys with and without attention-deficit/hyperactivity disorder. *J Consult Clin Psychol* 68:477–488, 2000.
- DuPaul GJ, McGoey KE, Eckert TL, VanBrakle J: Preschool children with attention-deficit/hyperactivity disorder: Impairments in behavioral, social, and school functioning. *J Am Acad Child Adolesc Psychiatry* 40:508–515, 2001.
- Eddy LS, Toro TJ, Salamero BM, Castro FJ, Cruz HM: Attention-deficit/hyperactivity disorder. A survey to evaluate risk factors, associated factors, and parental child-rearing behavior. *An Esp Pediatr* 50:145–150, 1999.
- Elia J, Welsh PA, Gullotta CS, Rapoport JL: Classroom academic performance: Improvement with both methylphenidate and dextroamphetamine in ADHD boys. *J Child Psychol Psychiatry* 34:785–804, 1993.
- Ernst M, Zametkin AJ, Matochik JA, Pascualvaca D, Jons PH, Cohen RM: High midbrain [¹⁸F]DOPA accumulation in children with attention-deficit/hyperactivity disorder. *Am J Psychiatry* 156:1209–1215, 1999.
- Evenden JL: Varieties of impulsivity. *Psychopharmacology (Berl)* 146:348–361, 1999.
- Faraone SV, Biederman J, Lehman BK, Spencer T, Norman D, Seidman LJ, Kraus I, Perrin J, Chen WJ, Tsuang MT: Intellectual performance and school failure in children with attention-deficit/hyperactivity disorder and in their siblings. *J Abnorm Psychol* 102:616–623, 1993.
- Fried I, Wilson CL, Morrow JW, Cameron KA, Behnke ED, Ackerson LC, Maidment NT: Increased dopamine release in the human amygdala during performance of cognitive tasks. *Nat Neurosci* 4:201–206, 2001.
- Friston KJ, Tononi G, Reeke GN, Jr, Sporns O, Edelman GM: Value-dependent selection in the brain: Simulation in a synthetic neural model. *Neuroscience* 59:229–243, 1994.
- Grodzinsky GM, Barkley RA: Predictive power of frontal lobe tests in the diagnosis of attention-deficit/hyperactivity disorder. *Clin Neuropsychol* 13:12–21, 1999.
- Haenlein M, Caul WF: Attention-deficit disorder with hyperactivity: A specific hypothesis of reward dysfunction. *J Am Acad Child Adolesc Psychiatry* 26:356–362, 1987.
- Heiligenstein E, Guenther G, Levy A, Savino F, Fulwiler J: Psychological and academic functioning in college students with attention-deficit/hyperactivity disorder. *J Am Coll Health* 47:181–185, 1999.
- Holland JH: Escaping brittleness: The possibilities of general-purpose learning algorithms applied to parallel rule-based systems. In: *Machine Learning II*. Edited by Michalski RS, Carbonell JG, Mitchell TM. Los Altos (California), Morgan Kaufmann, 1986, pp 593–623.
- Holland PC, Gallagher M: Amygdala circuitry in attentional and representational processes. *Trends Cogn Sci* 3:65–73, 1999.
- Horvitz JC: Mesolimbocortical and nigrostriatal dopamine responses to salient nonreward events. *Neuroscience* 96:651–656, 2000.
- Kacelnik A: Normative and descriptive models of decision making: Time discounting and risk sensitivity. *Ciba Found Symp* 208:51–67, 1997.
- Kakade S, Dayan P: Dopamine: Generalization and bonuses. *Neural Netw* 15:549–559, 2002.
- Kirby KN, Herrnstein RJ: Preference reversals due to myopic discounting of delayed reward. *Psychol Sci* 6:83–89, 1995.
- Kirby KN, Petry NM, Bickel WK: Heroin addicts have higher discount rates for delayed rewards

- than non-drug-using controls. *J Exp Psychol Gen* 128:78–87, 1999.
- Koob GF, Stinus L, Le Moal M, Bloom FE: Opponent process theory of motivation: Neurobiological evidence from studies of opiate dependence. *Neurosci Biobehav Rev* 13:135–140, 1989.
- LaHoste GJ, Swanson JM, Wigal SB, Glabe C, Wigal T, King N, Kennedy JL: Dopamine D4 receptor gene polymorphism is associated with attention-deficit/hyperactivity disorder. *Mol Psychiatry* 1: 121–124, 1996.
- Levy F, Hay DA, McStephen M, Wood C, Waldman I: Attention-deficit hyperactivity disorder: A category or a continuum? Genetic analysis of a large-scale twin study. *J Am Acad Child Adolesc Psychiatry* 36:737–744, 1997.
- Lockwood KA, Marcotte AC, Stern C: Differentiation of attention-deficit/hyperactivity disorder subtypes: Application of a neuropsychological model of attention. *J Clin Exp Neuropsychol* 23: 317–330, 2001.
- Madden GJ, Begotka AM, Raiff BR, Kastern LL: Delay discounting of real and hypothetical rewards. *Exp Clin Psychopharmacol* 11:139–145, 2003.
- Malhotra AK, Goldman D: The dopamine D(4) receptor gene and novelty seeking. *Am J Psychiatry* 157:1885–1886, 2000.
- Mannuzza S, Klein RG: Long-term prognosis in attention-deficit/hyperactivity disorder. *Child Adolesc Psychiatr Clin N Am* 9:711–726, 2000.
- McBride WJ, Murphy JM, Ikemoto S: Localization of brain reinforcement mechanisms: Intracranial self-administration and intracranial place-conditioning studies. *Behav Brain Res* 101:129–152, 1999.
- McClure SM, Berns GS, Montague PR: Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346, 2003.
- McLaren J: The development of selective and sustained attention in normal and attentionally disordered children. Dalhousie University, Halifax, Nova Scotia, 1989.
- Meneses A: Physiological, pathophysiological, and therapeutic roles of 5-HT systems in learning and memory. *Rev Neurosci* 9:275–289, 1998.
- Mirenowicz J, Schultz W: Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* 72(2):1024–1027, 1994.
- Montague PR, Dayan P, Sejnowski TJ: A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947, 1996.
- Monterosso J, Ainslie G: Beyond discounting: Possible experimental models of impulse control. *Psychopharmacology (Berl)* 146:339–347, 1999.
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ: Temporal difference models and reward-related learning in the human brain. *Neuron* 38: 329–337, 2003.
- O'Doherty JP, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ: Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304(5669):452–454, 2004.
- O'Reilly RC, Munakata Y: *Computational explorations in Cognitive Neuroscience*. Cambridge (Massachusetts), MIT Press, 2000.
- Oie M, Rund BR: Neuropsychological deficits in adolescent-onset schizophrenia compared with attention-deficit/hyperactivity disorder. *Am J Psychiatry* 156:1216–1222, 1999.
- Ownby RL, Harwood DG: Neuropsychological assessment of attention and its disorders: Computational models for neglect, extinction, and sustained attention. In: *Fundamentals of Neural Network Modeling: Neuropsychology and Cognitive Neuroscience*. Edited by Parks RW, Levine DS. Cambridge (Massachusetts), MIT Press, 1998, pp 257–269.
- Parasuraman R: *The Attentive Brain*. Cambridge (Massachusetts), MIT Press, 1998.
- Pashler HE: *The Psychology of Attention*. Cambridge (Massachusetts), MIT Press, 1998.
- Pineda D, Ardila A, Rosselli M, Arias BE, Henao GC, Gomez LF, Mejia SE, Miranda ML: Prevalence of attention-deficit/hyperactivity disorder symptoms in 4- to 17-year-old children in the general population. *J Abnorm Child Psychol* 27: 455–462, 1999.
- Popper CW: Antidepressants in the treatment of attention-deficit/hyperactivity disorder. *J Clin Psychiatry* 58(Suppl 14):14–29, 1997.
- Raggio DJ: Visuomotor perception in children with attention-deficit/hyperactivity disorder—combined type. *Percept Mot Skills* 88:448–450, 1999.
- Rappaport MD, Tucker SB, DuPaul GJ, Merlo M, Stoner G: Hyperactivity and frustration: The influence of control over and size of rewards in delaying gratification. *J Abnorm Child Psychol* 14: 191–204, 1986.
- Reed P, Mitchell C, Nokes T: Intrinsic reinforcing properties of putatively neutral stimuli in an instrumental two-lever discrimination task. *Anim Learn Behav* 24:38–45, 1996.
- Richards JB, Sabol KE, de Wit H: Effects of methamphetamine on the adjusting amount procedure, a model of impulsive behavior in rats. *Psychopharmacology (Berl)* 146:432–439, 1999.
- Rubia K, Overmeyer S, Taylor E, Brammer M, Williams SC, Simmons A, Bullmore ET: Hypofrontality in attention-deficit/hyperactivity disorder during higher-order motor control: A study with functional MRI. *Am J Psychiatry* 156:891–896, 1999.
- Rubia K, Taylor E, Smith AB, Oksanen H, Overmeyer S, Newman S, Oksanen H: Neuropsychological analyses of impulsiveness in childhood hyperactivity. *Br J Psychiatry* 179:138–143, 2001.

- Rutter M, Plomin R: Opportunities for psychiatry from genetic findings. *Br J Psychiatry* 171:209–219, 1997.
- Rutter M, Rutter M: *Developing Minds: Challenge and Continuity Across the Lifespan*. New York, Harper Collins, 1993.
- Sagvolden T: The attention-deficit disorder might be a reinforcement deficit disorder. In: *Contemporary Psychology in Europe: Theory, Research, and Application*. Edited by Georgas J, Manthouli M, Besevegis E, Kokkevi A. Göttingen, Hogrefe and Huber, 1996, pp 131–143.
- Sagvolden T, Aase H, Johansen E: The neuropsychology of attention-deficit/hyperactivity disorder (AD/HD) in a hypodopaminergic system. 6th Internet World Congress for Biomedical Sciences, INABIS, Ciudad Real, Spain, February 14–25, 2000.
- Sagvolden T, Aase H, Zeiner P, Berger DF: Altered reinforcement mechanisms in attention-deficit/hyperactivity disorder. *Behav Brain Res* 94:61–71, 1998.
- Sagvolden T, Sergeant JA: Attention-deficit/hyperactivity disorder—from brain dysfunctions to behaviour. *Behav Brain Res* 94:1–10, 1998.
- Scahill L, Schwab-Stone M: Epidemiology of ADHD in school-age children. *Child Adolesc Psychiatr Clin N Am* 9:541–555, 2000.
- Schultz W: Predictive reward signal of dopamine neurons. *J Neurophysiol* 80(1):1–27, 1998.
- Schultz W, Dayan P, Montague PR: A neural substrate of prediction and reward. *Science* 275:1593–1599, 1997.
- Servan-Schreiber D, Printz H, Cohen JD: A network model of catecholamine effects: Gain, signal-to-noise ratio, and behavior. *Science* 249:892–895, 1990.
- Shallice T, Burgess P: The domain of supervisory processes and temporal organization of behavior. *Philos Trans R Soc Lond B Biol Sci* 351:1405–1411, 1996.
- Shallice T, Marzocchi GM, Coser S, Del Savio M, Meuter RF, Rumiati RI: Executive function profile of children with attention-deficit/hyperactivity disorder. *Dev Neuropsychol* 21:43–71, 2002.
- Shekim WO, Dekirmenjian H, Chapel JL: Urinary MHPG excretion in minimal brain dysfunction and its modification by d-amphetamine. *Am J Psychiatry* 136:667–671, 1979.
- Slusarek M, Velling S, Bunk D, Eggers C: Motivational effects on inhibitory control in children with ADHD. *J Am Acad Child Adolesc Psychiatry* 40:355–363, 2001.
- Sonuga-Barke EJ, Saxton T, Hall M: The role of interval underestimation in hyperactive children's failure to suppress responses over time. *Behav Brain Res* 94:45–50, 1998.
- Sonuga-Barke EJ, Taylor E, Sembi S, Smith J: Hyperactivity and delay aversion—I. The effect of delay on choice. *J Child Psychol Psychiatry* 33:387–398, 1992.
- Spielewoy C, Roubert C, Hamon M, Nosten-Bertrand M, Betancur C, Giros B: Behavioral disturbances associated with hyperdopaminergia in dopamine-transporter knockout mice. *Behav Pharmacol* 11:279–290, 2000.
- Sunohara GA, Malone MA, Rovet J, Humphries T, Roberts W, Taylor MJ: Effect of methylphenidate on attention in children with attention-deficit/hyperactivity disorder (ADHD): ERP evidence. *Neuropsychopharmacology* 21:218–228, 1999.
- Sutton RS, Barto AG: *Reinforcement Learning*. Cambridge (Massachusetts), MIT Press, 1998.
- Swanson J, Oosterlaan J, Murias M, Schuck S, Flodman P, Spence MA, Wasdell M, Ding Y, Chi HC, Smith M, Mann M, Carlson C, Kennedy JL, Sergeant JA, Leung P, Zhang YP, Sadeh A, Chen C, Whalen CK, Babb KA, Moyzis R, Posner MI: Attention-deficit/hyperactivity disorder children with a 7-repeat allele of the dopamine receptor D4 gene have extreme behavior but normal performance on critical neuropsychological tests of attention. *Proc Natl Acad Sci U S A* 97:4754–4759, 2000a.
- Swanson JM, Flodman P, Kennedy J, Spence MA, Moyzis R, Schuck S, Murias M, Moriarity J, Barr C, Smith M, Posner M: Dopamine genes and ADHD. *Neurosci Biobehav Rev* 24:21–25, 2000b.
- Tannock R: Attention-deficit/hyperactivity disorder: Advances in cognitive, neurobiological, and genetic research. *J Child Psychol Psychiatry* 39:65–99, 1998.
- Tannock R, Schachar R: Methylphenidate and cognitive perseveration in hyperactive children. *J Child Psychol Psychiatry* 33:1217–1228, 1992.
- Tannock R, Schachar RJ, Carr RP, Logan GD: Dose-response effects of methylphenidate on academic performance and overt behavior in hyperactive children. *Pediatrics* 84:648–657, 1989.
- Taylor E: Syndromes of attention deficit and overactivity. In: *Child and Adolescent Psychiatry. Modern Approaches*. Third Edition. Edited by Rutter M, Taylor E, Hersov L. Oxford, Blackwell Science Ltd., Osney Mead, 1994, pp 285–307.
- Thapar A, Holmes J, Poulton K, Harrington R: Genetic basis of attention-deficit and hyperactivity. *Br J Psychiatry* 174:105–111, 1999.
- Usher M, Cohen JD, Servan-Schreiber D, Rajkowski J, Aston-Jones G: The role of locus coeruleus in the regulation of cognitive performance. *Science* 283:549–554, 1999.
- Waldman ID, Rowe DC, Abramowitz A, Kozel ST, Mohr JH, Sherman SL, Cleveland HH, Sanders ML, Gard JM, Stever C: Association and linkage of the dopamine transporter gene and attention-deficit/hyperactivity disorder in children: Heterogeneity owing to diagnostic subtype and severity. *Am J Hum Genet* 63:1767–1776, 1998.

- Walker AJ, Shores EA, Trollor JN, Lee T, Sachdev PS: Neuropsychological functioning of adults with attention-deficit/hyperactivity disorder. *J Clin Exp Neuropsychol* 22:115–124, 2000.
- Watkins C: Learning from Delayed Rewards. University of Cambridge, 1989.
- Weglage J, Pietsch M, Funders B, Koch HG, Ullrich K: Deficits in selective and sustained attention processes in early treated children with phenylketonuria—result of impaired frontal lobe functions? *Eur J Pediatr* 155:200–204, 1996.
- Wilkison PC, Kircher JC, McMahon WM, Sloane HN: Effects of methylphenidate on reward strength in boys with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 34: 897–901, 1995.
- Williams J, Taylor E: Dopamine appetite and cognitive impairment in attention-deficit/hyperactivity disorder. *Neural Plas* 11(1*2), 2004.
- Young SE, Stallings MC, Corley RP, Krauter KS, Hewitt JK: Genetic and environmental influences on behavioral disinhibition. *Am J Med Genet* 96: 684–695, 2000.
- Yu AJ, Dayan P: Acetylcholine in cortical inference. *Neural Netw* 15:719–730, 2002a.
- Yu AJ, Dayan P: Expected and unexpected uncertainty. ACh and NE in the neocortex. *Neural Information Processing Systems*, 2002b.

Address reprint requests to:

Jonathan Williams
Department of Child and Adolescent Psychiatry
Institute of Psychiatry
De Crespigny Park
Denmark Hill
London SE5 8AF
United Kingdom

E-mail: johwilliams@gmail.com

APPENDIX

In temporal-difference learning, using exponential discounting, the *discounted sum of future reinforcements* $V(t)$ should be:

$$V(t) = \left\langle \sum_{\tau \geq t} D^{\tau-t} r(\tau) \right\rangle$$

where t is the current time in the trial, τ is the future time, $r(\tau)$ is the reinforcement delivered at time τ , and D is the discount factor. Angle brackets $\langle \rangle$ indicate that this is averaged over the random choice of actions and the random reinforcements (e.g., the random termination of trials).

Montague et al. (1996) used exactly $V(t)$ as the critic in a form of the actor-critic architecture (Barto et al. 1983). In this paper, we use $V(t) = Q(t, u)$, where $Q(t, u)$ (whose values are presented in the tables in the main text) depend (Watkins 1989) on the action u selected by the model at time t , where u is either *act* or *wait*. The model gradually comes to estimate the correct value of $Q(t, u)$ for every explored time-action combination by adjusting its estimates based on prediction errors:

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

where $r(t)$ is the actual reinforcement obtained at time t . The adjustment to the parameter $Q(t, u)$ is

$$Q(t, u) \rightarrow Q(t, u) + \eta \delta(t)$$

if action u is actually taken at time t , whence also $V(t) = Q(t, u)$. Here, η is the learning rate.

We have, so far, discussed how temporal-difference learning comes to make predictions about future rewards. How can subjects use these predictions to choose among possible actions? Temporal-difference learning has explored a number of different possibilities. We adopted a simple one in our study, in which the subjects learned a different prediction $Q(t, u)$ for each possible action (u ; *act* or *wait*) at each timestep (t) (Watkins 1989). At each timestep the model makes a decision about whether to act or not (i.e., whether to press the lever) based on the expected sum of future reinforcements in either case. Some randomness was added into this behavior, to permit exploration, using the softmax function:

$$p_t(u) = \frac{e^{\beta(Q(t, u) + \phi_u)}}{\sum_{u'} e^{\beta(Q(t, u') + \phi_{u'})}}$$

where $p_t(u)$ is the probability of performing a particular action u , ϕ_u is a fixed bias for particular actions, and β is the brittleness. Action bias could, in principle, be moved out of the parentheses, changing the interaction with brittleness (broadly) from multiplication to addition, but there is insufficient experimental data to be certain that either way is better.

For Figure 6, we assumed that, as a child grows up, his discount factor D increases, so that delayed reinforcements come to seem more valuable. At each timestep, D was adjusted toward $D' = 0$ or 1 , depending on whether or not the trial was interrupted at that timestep:

$$D_{t+1} = D_t + \eta_m (D' - D_t)$$

where η_m is the learning rate for the metalearning. We set $\eta_m = 0.0002$, making it much slower than trial learning.

In Figure 4, we calculated the entropy of the timestep at which the lever is pressed, using Shannon's definition for entropy, as follows:

$$H = - \sum_{0 < t \leq 10} act_t \log_2 act_t$$

where act_t is the proportion of trials on which the lever was pressed at timestep t . This formula gives the average number of bits needed to encode behavior in a trial. (So, for example, if the subject always pressed the lever at exactly the same timestep, there would be no entropy, and $H = 0$.)