Dopamine Appetite and Cognitive Impairment in Attention Deficit/Hyperactivity Disorder

Jonathan Williams[1,2] and Eric Taylor[3]

Department of Child and Adolescent Psychiatry, Institute of Psychiatry
De Crespigny Park, Denmark Hill, London SE5 8AF, UK

SUMMARY

The underlying defects in ADHD (Attention Deficit/Hyperactivity Disorder) are not yet clear. The current paper tests three existing theories: State Regulation, Cognitive Deficit, and Temporal Difference (TD) learning. We present computational simulations of the Matching Familiar Figures Task and compare these with the experimental results reported by Sonuga-Barke (2002). The TD model contains four parameters: the learning rate, discounting for future rewards, brittleness (randomness) of behavior, and action bias. The results show that the basic TD model accounts well for control performance in trials of 5 sec, 10 sec, and 15 sec duration; but not for the deficits in ADHD performance at 5 sec and 15 sec. Extending the TD model to incorporate either a state regulation deficit, or working memory deficit and delay in starting trials, can provide a good account of both control and ADHD results, at all trial-lengths. We discuss the significance of the results for theories of ADHD and make suggestions for future experimentation.

KEYWORDS

dual pathway, methylphenidate, computational cognitive deficit, state regulation, ADHD

INTRODUCTION

The Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) (American Psychiatric Association, 2000) offers three subtypes of attention deficit hyperactivity disorder (ADHD): Predominantly inattentive; Predominantly hyper-active-impulsive; and Combined, implying that there are two main causes of ADHD, which can co-occur. A similar view has been codified as the "dual pathway model" (Sonuga-Barke, 2002b). The dual causation view has support from several sources. O ne is the finding that ADHD children with a susceptibility conferring 7-repeat allele of the dopamine D4 receptor gene are remarkably free of the slow, highly variable response times that characterize ADHD children lacking that allele (Swanson et al., 2000a). Furthermore, two quite different tasks, the Choice-Delay Task (CDT) and the Stop Signal Task (SST) together form a very sensitive and specific predictor of the clinical diagnosis of ADHD. Yet the two measures show very little correlation between themselves (Solanto et al., 2001), supporting the independence of the two causative pathways. The SST is seen as revealing a component of executive control, which may be related to general cognitive level; and the

Reprint requests to: Jonathan Williams, Department of Child and Adolescent Psychiatry, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, UK; e-mail: johwilliams@hotmail.com
Financial support received from: [1]London Postgraduate Medical Deanery; [2]Gatsby Charitable Foundation; [3]Medical Research Council.

© 2004 Freund & Pettman, U.K.

CDT appears to measure delay aversion, which may be caused by an abnormality of reward processing.

Although several etiological factors have been associated with ADHD, the mechanisms by which they act are still largely mysterious. However, a new level of theoretical clarity was achieved recently with the head-to-head testing of four explanatory theories of ADHD, using the Matching Familiar Figures Task (MFFT), in both a severely affected clinical sample and a mild population sample matching criteria for combined type ADHD (Sonuga-Barke, 2002a). In this task, seven pictures are presented simultaneously on a computer screen, and the subject has a specified number of seconds to discover which picture matches the top one. Each of the four theories (State Regulation Deficit, (General) Cognitive Deficit, Premature Task Disengagement, and Ecological Niche) led to distinct predictions regarding ADHD performance on the MFFT.

In short, the control groups' performance steadily improved as the trials were lengthened from 5 to 10 then to 15 seconds. In comparison, the ADHD performance was inferior to that of the controls at 5 sec; then equivalent at 10 sec, and again inferior at 15 sec (shown below). This effect was seen in both of the very different samples, suggesting that we are seeing an important effect. The data will be used in the current paper as the basis for assessing the validity of three distinct computational simulations of decision-making in ADHD, which can be briefly described as learning-based, appetite-based, and cognition-based.

Cognitive deficits in ADHD

It is well known that children with ADHD often suffer from academic impairment (Barkley et al., 1991; Faraone et al., 1993). Some impairments persist from preschool to college age (DuPaul et al., 2001; Heiligenstein et al., 1999).

Neuropsychological deficits have also 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). A 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 & Rund, 1999; Raggio, 1999; Sunohara et al., 1999).

Subgroups can be particularly affected (Swanson et al., 2000b). Perhaps half of all children with ADHD have significant deficits in motor or perceptual skills, without clear mental retardation or major neurological disability (Gillberg, 2003). Dyslexia and other language problems are seen in about 50% of children with ADHD (Gilger & Kaplan, 2001) and also often verbal-performance discrepancy on IQ testing.

It is not clear whether there is a visual search deficit in ADHD, but it seems likely (Mason et al., 2003; Malone & Swanson, 1993). More relevant to MFFT, however, children with ADHD are delayed in the initiation of a serial visual search and have deficits in spatial working memory (Karatekin & Asarnow, 1998; Kempton et al., 1999).

Difficulties after rewards can be expected for two reasons. First, reward may slow down ADHD children more than control children (Scheres et al., 2001). Also, set-shifting problems have been reported in preschool children with ADHD (Sonuga-Barke et al., 2003), and these may slow down the transition from processing reward to starting the next trial.

Dopamine appetite and state regulation in ADHD

It has been suggested that the underlying cause of ADHD is a defect in the regulation of activation: "To counteract a performance decrement, subjects have to regulate their state: they have to inhibit activation when stimuli are rapidly presented, and to excite activation when stimuli are slowly presented"

(van der Meere et al., 1999). The proposed regulation can be described as indirect because a predictor of need (the stimulus rate) is being used to regulate activation rather than the level of activation being used to regulate itself. In the control of respiration, where regulation has been well studied, such indirect regulation does exist, controlled by emotion, body position, movements, and predicted activity level (Shea, 1996); but direct control by hypercapnia is simpler and more important. We have therefore studied direct regulation as a potentially more parsimonious explanation for ADHD.

Stimulants like methylphenidate (MPH) block dopamine transporters (Volkow et al., 1998) and are used to treat ADHD. in which they improve academic performance and normalize various neuropsychological measures (Elia et al., 1993; Haenlein & Caul, 1987; Sunohara et al., 1999; Tannock et al., 1989). Stimulants cause an immediate, unlearned, and fully reversible reduction in hyperactivity in ADHD and in control children (Hungund et al., 1979), leading to the possibility that the children have an 'appetite' for dopamine, which can be satisfied by stimulants.

Dopaminergic systems are activated by reward but also by various other factors, notably novelty (Ljungberg et al., 1992; Cloninger, 1987; Horvitz, 2000; Schultz, 1998). 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 & Goldman, 2000). Hyperactivity in ADHD has been repeatedly found to be ameliorated by novelty (Iaboni et al., 1995; Felicetti & Julliard, 2000; Sleator & Ullmann, 1981) and by high reinforcement rates (Douglas & Parry, 1994; Carlson & Tamm, 2000).

The Temporal Difference (TD) model of ADHD

This model is primarily concerned with the behavioral effects of variations in constraints (or parameters) controlling learning and behavior. The mapping of these parameters to brain structures (as in Fig. 1) is considered mainly where the experimental data is strongest, namely with the dopaminergic neuromodulatory system originating in the ventral tegmental area (VTA).

The link from the TD 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 substantia nigra of 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 activity of dopaminergic cells signals mismatches in these predictions. The model accounts well for a wide variety of data on the dopamine system in learning (Schultz et al., 1997), for which evidence is also accumulating in humans (Fried et al., 2001).

Figure 1(a) shows the basic TD model. In it, a representation of the current state is available in cortex. The basal ganglia learn to associate these states, and potential actions, with reinforcements. The 'action' (such as inspecting a picture or looking around the room) that appears likely to produce the largest reinforcement is in general chosen by the basal ganglia. The 'predicted reinforcement' is used in conjunction with information about immediate reinforcement to create the 'prediction error' signal that, in turn, is used to alter the prediction that will be made when, in future, the same situation is encountered again.

The model gradually learns to choose behaviors that lead to rewards. Details are given in the Appendix. The basic TD model used in this paper has already been validated against the Choice-Delay Task (CDT) (of Sonuga-Barke et al., 1992; Williams & Dayan, unpublished observations). However, a weakness of the model is that it does not provide an obvious account of the immediate effects of stimulants. The MFFT data are

particularly appropriate for testing the model, as it is quite different from and considerably more complicated than the CDT.

The current study presents a computational simulation of the task used by Sonuga-Barke (2002a), in order to determine whether state regulation deficits and/or cognitive impairment and/or the TD model are able to account for ADHD performance in the MFFT. Before starting this work, we expected, based on results described above, that the TD model needed to be extended to incorporate a State Regulation deficit.

EXPERIMENTAL

Modeling the MFFT

The challenge here is to simplify the task sufficiently to allow it to be programmed, without losing any aspect that is critical to ADHD. The task is implemented using a computer program. Time is simulated as steps, with each time-step notionally one second. At each time-step, the model has three options: to inspect the next of six target pictures to see whether it matches the index picture; to similarly inspect a randomly chosen picture; or to do something else (not specified).

The purpose of the inspections is for the model to steadily improve its estimation of the likelihood that each target picture is a match for the index picture. After all the time-steps in the trial have elapsed, the model is interrogated regarding its choice of best match. It then has two options: to give its best guess or just to give a random answer.

Complicated cognitive tasks, such as task acquisition and visual search, are not simulated in this work. Visual search in particular has been the subject of a large amount of theoretical and simulation work, but none has directly addressed the MFFT.

The model treats quite differently the situations of inspecting a target picture that matches the index, versus inspecting a non-match. We consider the latter first. When choices are superficially similar (i.e. in difficult trials), recognizing differences is more likely than recognizing the identical match. So the program keeps track of the current estimated probability for each of the six target pictures. These start at 1/6. When a difference is found, the probability of that picture matching the index drops toward zero, at a rate determined by the 'learning rate', and the probabilities are re-distributed to total 1.

However with very simple pictures (e.g. stick figures), subjects may be able to tell 'at a glance' that they have found the right one. Our simulation allows such recognition in easy trials only. Without such an 'at a glance' method, the model cannot achieve learning as fast as the experimental controls (though such fast learning could also be accounted for by controls having a special quick-scan strategy for short trials.) Note that the distribution of difficulty levels of the trials affects performance (see Appendix).

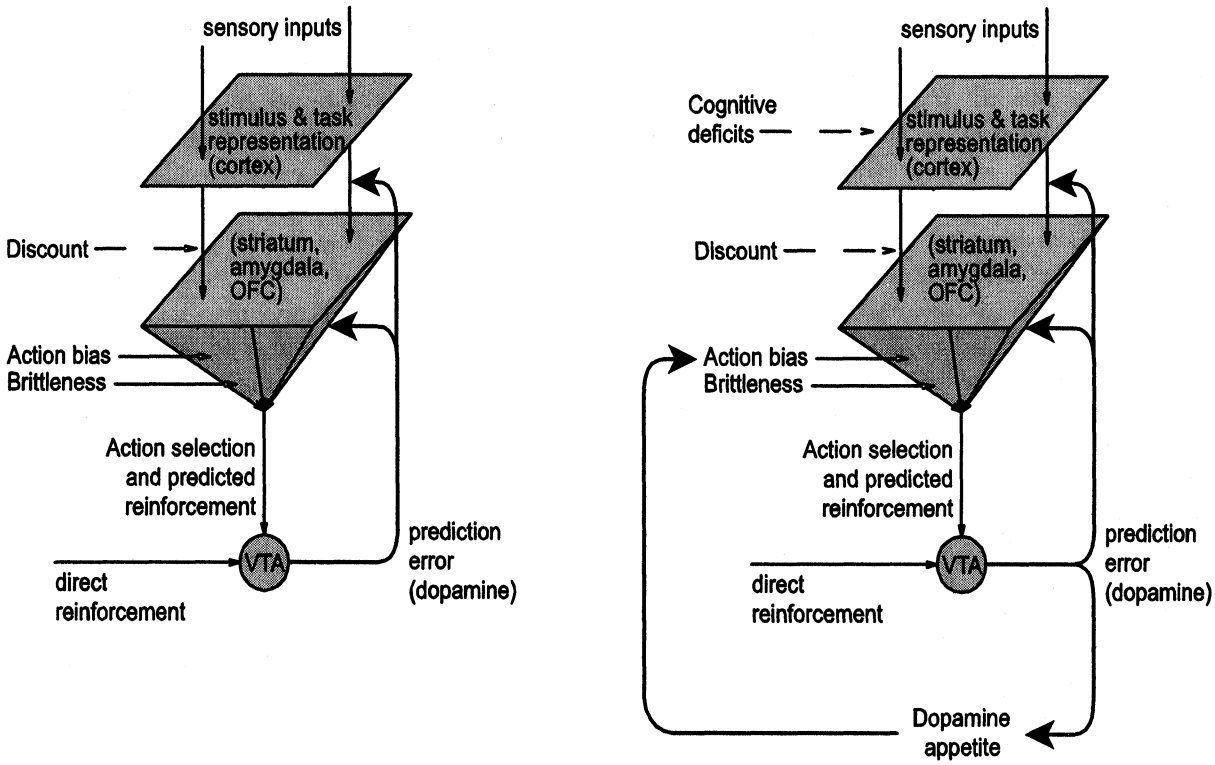
The overall approach just described can be implemented in many ways. Our program incorporated temporal difference learning, with its parameters described below. It also incorporated parameters suggested by the state regulation and cognitive deficit theories (also described below), in such a way that any or all of them could be disabled, and the effects of the remaining one or two factors could be investigated in isolation. The program is written in Matlab (available by email from the first author).

Applying the Basic Temporal Difference (TD) Model of learning to the MFFT

The TD model's behavior is controlled by the reward size and by four key parameters. In this context, parameters are numerical measures of long-term aspects of an organism's or the model's behavior. These are generally unchanging during an experimental episode and may be genetically

controlled. Although the parameters interact to determine overall behavior of the model, we study primarily the simple case of their effects in isolation.

1. *Perceived reward size.* This is conceived as the amount of pleasure elicited by a reward when it is received. In reality, it will be a nonlinear function of the objective reward size. We included this factor to allow the model to account for any child who tended to experience rewards as less rewarding than other children, as hypothesized as one possible cause of ADHD (Blum et al., 2000;



Model A
The Basic Temporal Difference Model

Model B
The Extended Temporal Difference Model. Differences from Model A are the inclusion of dopamine appetite and cognitive deficits.

Fig. 1: Models used in this study. (OFC: orbitofrontal cortex; VTA: Ventral Tegmental Area)

see discussion in Solanto et al., 2001). The explicit reward provided by the experimenter at the end of successful trials was modeled. We did not model, however, implicit rewards in the task, such as the novelty of seeing new pictures, the pleasure of finding each difference, and the pleasure in becoming certain of a match.

2. *Reward discounting.* The idea here is that a reward expected in the future is worth less than the same reward delivered now. Repeatedly in life, individuals face choices between small, immediate rewards and delayed large rewards (e.g. in dieting). We quantify this by multiplying the reward by a number D (between 0 and 1), at each time-step. At 9 pm, an immediate reward might be worth R, but at 8 pm, it is worth RD, and at 7 pm, it is worth even less, RD^2. In our model, ADHD children have a smaller D, so future reinforcements are discounted more, and delayed rewards will influence their behavior less than in control children. Indeed, increased discounting has been demonstrated in ADHD (Barkley et al., 2001; Sagvolden et al., 1998); Sagvolden, Aase et al (2000) proposed that a "shorter and steeper delay gradient" in ADHD children may lead to the development of overactivity, increased behavioral variability, motor impulsiveness, and impaired sustained attention.

Such a gradient has also been correlated with impulsive behavior in general psychiatric outpatients (Crean et al., 2000). Males discount future rewards more than females, and addicts more than non-addicts (Wilson & Daly, 2003; Kollins, 2003). It has been suggested (Williams & Dayan, 2004) that the discounting rate may to some extent be learned from the degree of unpredictability in an individual's environment; such unpredictability, however, may be highly context-specific (Aloise & Miller, 1991). O ver a much shorter time course, discounting can be increased (i.e. D can be reduced) by exposure to certain pleasurable stimuli (Wilson & Daly, 2003).

3. *Brittleness.* The predictability of the model's behavior is dependent on three main factors: the reliability of environmental predictors of reinforcement, the extent to which such lessons have been learned, and finally brittleness, the extent to which behavior is based on those learned lessons. Brittleness being too high allows single learning experiences to cause sudden changes from one consistent behavior to another, and a comparative inability to persist with one behavior in an unpredictable environment. On the other hand, brittleness being too low would indicate a comparative inability to persist with one behavior, even when the subject has collected adequate information about reinforcement availability.

In principle, a subject might be more predictable in some areas of decision-making than in others. The first is the brittleness of the decision-making at every time-step. Here the subject must choose whether to (a) methodically inspect the next picture in sequence; (b) randomly choose another picture to inspect; or (c) do something else. The second is in communication of the choice at the end of each trial, when the subject must choose whether to (a) tell the interviewer the most likely match; or (b) guess.

4. *Learning Rate.* For consistency with the computational literature, this term is taken to mean the maximal rate at which the model is able to alter its predictions about the environment, rather than, as would be more natural from a behaviorist perspective, the rate at which behavior changes. In the model, prediction error causes changes in predictions and future actions. If the error signal is multiplied or divided by a factor, then this has the same effect as changing the response to that signal, i.e. the learning rate. Although we model the learning rate as primarily dependent on various heritable factors associated with dopamine, the rate may also reflect intrinsic synaptic factors. Therefore, the terms 'dopamine multiplier' and 'learning rate' are interchangeable.

There are several lessons learned by subjects in this task, including (a) the rules of the task; (b) strategies to use; (c) within an individual trial, the characteristics of the index picture; and (d) which pictures are different from the index picture. Both (a) and (b) were not needed to explain the data from Sonuga-Barke because the order of trial lengths was randomly determined for each subject. For simplicity, we have excluded (c) and modeled only (d).

A high learning rate is a disadvantage in tasks for which previous learning can excessively impair exploration. Such is the case with the Delayed Response Time Task (DRTT; Williams & Dayan, 2004) and in exploring an environment to find maximal reward. T his would be relevant to the MFFT in trials in which a subject erroneously identified a match, and so then stopped exploring and did not have a chance to correct his mistake. This sort of mistake has not been included in the current simulation.

5. *Action Bias.* Action bias is a measure of the child's preference for action over inaction. Action bias is greater than zero if the child, for any reason, has an innate bias to act, such as finding that action itself reinforcing. Co nversely, action bias is negative if the child prefers to be inactive. A non-zero action bias can force the child to make suboptimal decisions. Such a preference could in theory be innate or learned or both.

There is little opportunity within the MFFT task for physical action, and so we have chosen to define action bias as the tendency to prefer activity away from the task.

Extensions to the Temporal Difference Model used in Model B

The TD model was extended to incorporate state regulation problems and cognitive impairment to see whether these allow a better fit to the experimental results. These extensions are shown in Fig. 1 (b) and are discussed below.

State regulation deficit

We have included dopamine appetite as a simple example of state regulation. This takes the form of a 'dopamine level' that is increased when reward is received at the end of trials, and then slowly decays away.

The dopamine appetite must be linked to some output to complete a regulatory loop. In principle, appetite could act by increasing or decreasing any other parameter (or parameters) in the model. The 'gearing' of this influence is a numeric value that we can adjust in the simulation. We have simulated appetite effects on action bias and brittleness, because these parameters can have immediate and reversible effects on behavior (see Fig. 1), as suggested by the immediate result of stimulants. We have not simulated appetite effects on learning rate or discount, because the model does not allow these to have immediate reversible effects.

Cognitive impairment

Performance of the MFFT involves many subtasks. We have modeled the effect of several changes that might be expected to accompany general cognitive impairment. Each is potentially very complicated, so we have not modeled them in detail. Instead, we have represented each by a single number.

1. *Working memory.* ADHD is often accompanied by an impairment in working memory (Sonuga-Barke et al., 2003). In real subjects, working memory is important for holding features of the index picture in mind, for comparison with the others. Working memory is also important for holding the 'result' until it is requested by the experimenter. We have modeled only this latter aspect. We model working memory as holding estimates of the likelihood of each picture matching the index picture. These estimates are subject to slight 'blurring' during each second in which the model is off-task (see Appendix for details).

We expect that most children are motivated both by their desire to succeed and by their desire to please the investigator. After the children become certain that they have found the right answer, the question of succeeding disappears, and they may become interested in something else. We model this by stopping all activity on the task, once an exact match has been found. In long trials where the model finds a match early, the subsequent fading of its working memory may limit performance.

2. *Delay at the start of trials.* This delay is the time used between finishing one trial and starting the search through the object pictures. The delay consists of the Post-Reinforcer Pause (PRP) plus the time used in any initial scan of the index picture and will include the time needed for set-shifting, which has been shown to be an area of deficiency in ADHD (Sonuga-Barke et al., 2003). During this time, the model is losing opportunities to collect information about the pictures to match.

Other cognitive factors

a) *The ability to search from picture to picture methodically, avoiding time-wasting repetition.* This ability helps performance in short trials but not in long trials; in any event, it cannot account for an absolute decline in performance in longer trials (results not shown).
b) *The ability to find any difference between the index picture and the picture being inspected.* This ability helps performance at all trial lengths, but like (a) cannot account for an absolute decline in performance in longer trials, so is not shown.
c) *Speed of thinking.* Many studies have shown apparently slowed thinking/responses in ADHD (e.g. Mason et al., 2003; Solanto et al., 2001; Berman et al., 1999). We have not modeled this in detail because it cannot provide a parsimonious account of the MFFT differences between ADHD and controls. Although stretching of time can be used to account for the difference at one trial-length, extra factors would be required for the other two trial-lengths.

RESULTS

Model A

1. *Perceived reward size.* A reduction in perceived reward size impairs performance across a wide range of trial-lengths (results not shown but very similar to Figure 2(a)). This occurs because, in the model, a substantial reward increases the likelihood of (i) at each time-step, staying on-task and exploring the pictures carefully; and (ii) at the end of each trial, telling the investigator the correct answer.

2. *Reward discounting.* Increased discounting has minimal effect in short trials; impairs performance somewhat in longer trials; and has little effect in very long trials (result not shown). The deterioration is caused by a reduction in incentive, during the early part of long trials, for the individual to work at the task. In very long trials, this wasted time does not matter. All results shown use no discounting (i.e. D = 1).

3. *Brittleness.* Groups with low brittleness of communication have little tendency to tell the interviewer what they know (this would be quite different from having an oppositional inclination to tell him a falsehood). Their recorded performance is very poor and is little affected by what they may have learned (results not shown). Groups with high brittleness of communication, on the other hand, reliably tell the interviewer what they know. This allows the interviewer to obtain a good measure of their ability. All results shown use a high level of brittleness (6).

4. *Learning Rate.* A higher learning rate is generally beneficial in this simulation, as we have not modeled erroneous identification of matches. The exception is in cases for which it leads to premature task disengagement (results not shown;

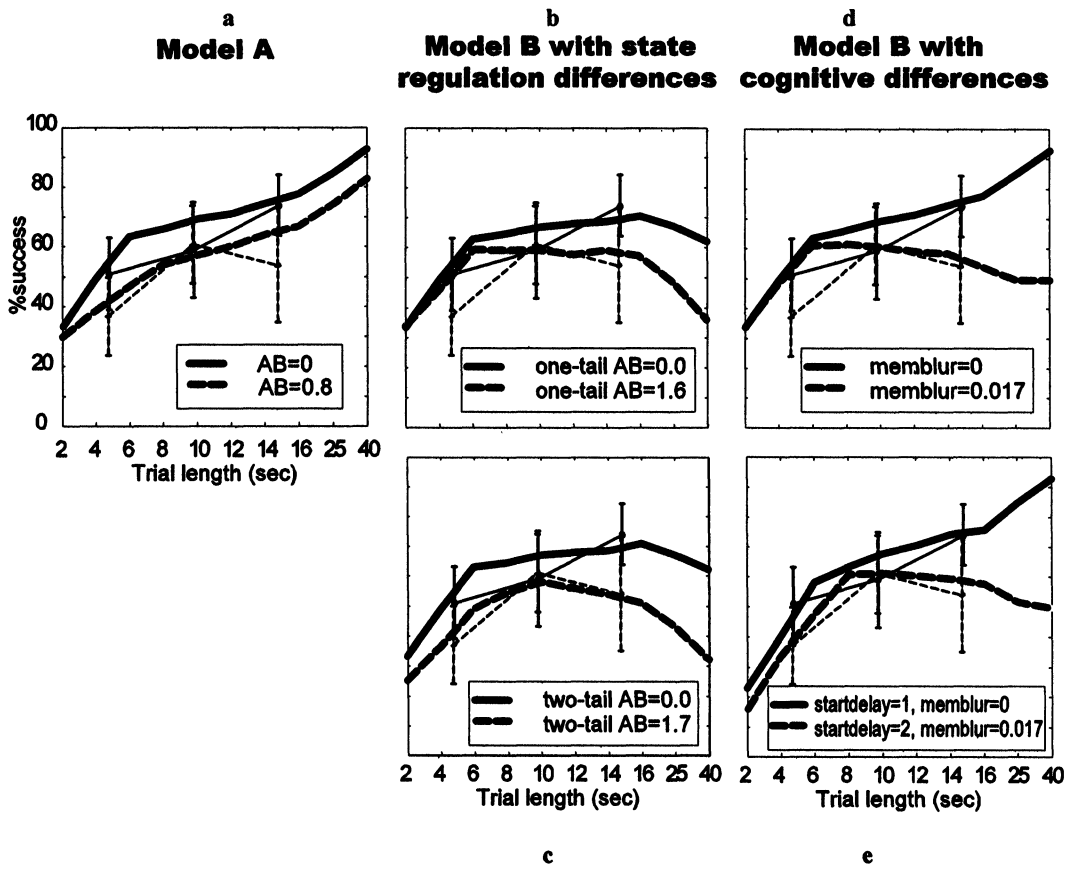


Fig. 2: Simulations of the Matching Familiar Figures Task (MFFT), compared with real data. Each graph shows the means and standard deviations from Sonuga-Barke (2002) Study 1, redrawn unchanged in every graph as thin lines (solid=control; dashed=ADHD). Superimposed on this, each graph shows the performance of two simulated groups of children, as thick lines (solid=model of controls; dashed=model of ADHD). All the simulations use the same values for all parameters, except for the one or two explicitly manipulated for each figure and shown in the legend boxes, and Model A does not include any working memory deficit, delay at the start of trials, or state regulation mechanism. Each simulated data point results from 5000 trials. Note that the right ends of the x-axes are compressed. (a) Effect on the model of fixed difference in Action Bias. (b) Effect on the model, of fixed difference in response to reward deficit, increasing Action Bias. (c) Effect on the model, of fixed difference in response to nonoptimal reward rate, increasing Action Bias. (d) Effect on the model, of reducing the retentiveness of working memory. (e) Effect on the model, of two fixed differences: start-trial delay and working memory retentiveness.

Abbreviations: AB = Action Bias (see text). One-tail AB: The extent to which non-optimal reward rate increases the Action Bias. This is "one-tailed" because reduction, but not increase in the reward rate increases Action Bias. Two-tail AB: The extent to which non-optimal reward rate increases the Action Bias. This is "two-tailed" because increase in the reward rate increases Action Bias, just as reduction does. Memblur: the extent of blurring of working memory each second (see Appendix). Startdelay: The number of seconds not used at the start of each trial.

see Discussion). All results shown use a high learning rate (1).

5. *Action bias.* Figure 2(a) is shown as an example of the behavior of Model A, in the Matching Familiar Figures Task. The figure shows the effect of increasing the action bias on the performance of the model. This impairs performance on the task, as time away from the task reduces time spent inspecting the pictures.

Summary of results from Model A

Learning creates a general trend to better performance in longer trials, seen in the simulations. This effect succeeds in very roughly approximating the improvement seen in real ADHD and control groups (see Fig. 2(a)). The parameters of the TD model can provide various explanations for the tendency for ADHD performance to be somewhat worse than the control group. These explanations include reductions in perceived reward size, brittleness, or learning rate, and increases in reward discounting or action bias.

However, as shown for example in Fig. 2(a), simple changes to the TD parameters are not able to account for the absolute deterioration in ADHD performance seen empirically in long trials.

Model B, with a Difference in State Regulation between the two groups

We introduced a link from deficient reward (i.e. below the set-point) to an increased Action Bias. In practical terms, this reflects the tendency of a child to become more active when he has not received much stimulation recently. The effect of this factor is to selectively impair performance on longer trials, as expected (Fig. 2(b)).

Figure 2(c) shows a two-tailed version of the same mechanism. In this case, Action Bias is increased whenever the reward rate (modeled as dopamine level) is outside an acceptable range. This is intended to model previous suggestions that children could have their performance impaired both by becoming over-excited at high rates and disinterested at low rates.

The possibility of a link from reward excess or deprivation to reduced brittleness (rather than to action bias) was also briefly explored. In the model, such a link was much less impairing than the link to increased Action Bias (result not shown).

Model B, with difference in Cognitive Impairment between the two groups

1. *Working memory.* Figure 2(d) shows that impaired retentiveness of working memory has a small effect in short trials, but a major effect in long trials. This is because after the model has reached its criteria for success in a trial, it tends to choose non-task activities, during which its working memory fades. Other simulations (not shown) showed that having a lower threshold for certainty, i.e. disengaging from the task before the answer was well known, preferentially impaired performance on long trials, as expected.

2. *Delay at start of trials.* Figure 2(e) shows the effect of hypothesizing two related differences between ADHD and control: working memory retentiveness and start-trial delay. The upper solid line indicates the performance of the model when not losing any such search time. The lower solid line indicates the performance of the model when 3 seconds are lost at the start of every trial. The figure shows that the effect of such disruption reduces as trial length is increased, as one would expect.

DISCUSSION

Main results

We have attempted to explain the difference between ADHD and control performances on a high-level cognitive task, the MFFT, using the Temporal Difference (TD) model of ADHD. The

TD model is a biologically based multifactorial theory explained in detail elsewhere (Williams & Dayan, 2004). We have shown that the TD model of ADHD is able to provide a fair account of the performance of both ADHD and control groups on the MFFT. The model can account for impairment in ADHD performance in long trials, relative to controls [Fig. 2(a)] but not the absolute impairment repeatedly reported (van der Meere et al., 1999; Dalby et al., 1977; Sonuga-Barke, 2002a). The model can also account for impaired ADHD performance in short trials (5 sec) but not for this impairment being greater than that seen at 10 sec.

We have therefore tested two extensions to the TD model. The main finding was that both the State Regulation theory and the Cognitive Deficit theory are individually able to account for the short-trial and long-trial deficits in MFFT performance seen in ADHD children relative to controls (Figs. 2(c) and 2(e), respectively). T he model also demonstrated interactions between these two mechanisms: cognitive abilities created conditions (such as lack of reward) in which state regulation mechanisms dominated behavioral choices; conversely, state regulation mechanisms (such as the proposed link from reward deficiency to action bias) could prevent adequate cognitive exploration of the pictures in the trial.

Despite the similarities between (c) and (e) in Fig. 2, the most straightforward explanations for short- and long-trial deficits were provided by cognitive deficits and dopamine appetite, respectively. It thus seems likely that short-trial deficits will be preferentially correlated with execution function deficits; and long-trial deficits preferentially with delay aversion, stimulant efficacy, and dopaminergic genes. The need for two processes rather than one to explain the MFFT results is supported by the dissociation between response time and error rate in a GO-NO GO task (van der Meere et al., 1999). That study showed the error rate to be minimum at 4-sec inter-stimulus interval (ISI) and worse at 1-sec or 8-sec. The response time, however, was considerably shorter with the 1-sec ISI.

Our results are consistent with the truism that a single psychological test cannot in isolation distinguish among all the different causes of psychopathology. Any behavioral result has several potential causes. For example, using current data, we cannot separate the effects of brittleness from that of perceived reward size—or indeed, from oppositionality.

The computational approach described in this paper is biologically based, repeatable, and specified at a lower level than most psychological theories of ADHD. The model, in having numeric parameters, is consistent with a widely held view that the categorical DSM concept of 'ADHD' should be restated as continua. This aspect may be useful in attempts to relate behavioral changes seen in ADHD to underlying genetic and pharmacological influences.

A major strength of this study is that we have shown that the TD model, designed for a very different class of tasks (Choice-Delay Task and Delayed Reaction Time Task), can make a useful contribution to the interpretation of experimental results in much more complicated tasks, such as MFFT, because even complicated tasks involve separable aspects of learning and reward processing.

Relation to previous work

The simulations presented here partially support the conclusion of Sonuga-Barke (2002a) that a State Regulation deficit is the best single explanation for ADHD children's performance on the modified MFFT reported in their paper. The difference between ADHD and control children can be accounted for by a fixed difference in response to nonoptimal reward rate (Fig. 2(c)). This is the most parsimonious explanation we found, as a single parameter governs ADHD performance deficit in both long and short trials. The presence of such symmetrical activation of the

Action Bias is readily testable, but is an unnecessarily strong prediction: For example, state regulation difficulties may cause the long-trial deficit while cognitive deficits cause the short-trial deficit. And although children with Combined type ADHD were studied, it may well be that the short-trial and long-trial deficits were independently distributed.

A major attraction of the State Regulation view is that it provides a straightforward explanation of the immediate and reversible effects of stimulants and novelty. We made a specific interpretation of the State Regulation hypothesis, namely that humans attempt to regulate their exposure to dopamine, and that when their supply is inadequate they attempt to increase it by increasing their Action Bias. Th is interpretation is consistent with the MFFT results (Fig. 2(b-c)) and with the direct effects of reward and novelty in ADHD (references above).

Our simulations also indicate that broadly defined Cognitive Deficits can also explain the MFFT data. In particular, impairment specific to long trials may be caused by working memory deficits, and impairment specific to short trials may be caused by slowness in processing a reward and/or starting trials. In the model, impairment at all trial durations can be caused by difficulty in recognizing differences between pictures (result not shown), and such a cause might be expected from the differing mean mental ages of the control and ADHD groups used by Sonuga-Barke (2002a), namely 10.16 and 8.7 years.

Many children with ADHD are simultaneously afflicted in multiple ways, including mood disorders, tic disorders, and oppositionality (Jensen et al., 2001). A group of such children may give quite the wrong impression if they are assumed to have a single underlying pathology.

Dual-Pathway Hypothesis

Our 'single-path' model involves the same transmitters, brain areas, and broadly defined executive dysfunctions (EDF) as the dual-pathway model of Sonuga-Barke (2002b). The models are complementary rather than competing, as the TD model focuses on the roles of dopamine and reward circuitry, whereas the dual-pathway model focuses on the roles of individual brain areas and describes compounding and compensatory processes in some detail.

The major conceptual difference between the two models is that whereas the elements of the TD model are working together on every task, the dual-pathway model has two paths making distinct contributions to behavior. Even though cortical areas are structurally and functionally distinct, however, the status of the suggestion that they form such well-segregated circuits (Alexander et al., 1986) is currently uncertain. Mesencephalic dopamine appears to be released throughout the striatum and frontal cortex as part of a unified wave, so it may not have the functional capacity to send distinct signals (Schultz, 1998). M oreover, parallel corticostriatal projections converge so dramatically in the striatum that any segregation of signals may be lost (Rolls & Treves, 1998).

Higher stimulant doses may have adverse cognitive effects while maintaining their anti-impulsivity effect (Tannock et al., 1995; Berman et al., 1999; Evans et al., 2001; O'Toole et al., 1997). This observation suggests at least two pharmacological effects, but whether the dissociation is between striatum and cortex, state-regulation and cognition, or even dopamine and norepinephrine, is not yet clear.

Limitations

The most important criticism of this work is that the small number of experimental data points could have been accounted for by an almost limitless range of physical processes. This problem is to some extent mitigated by our implementation of existing theories, rather than new ones invented for the purpose; by our use of fairly standard or

common-sense values for parameters; and by our adjusting only one of the parameters for each of the graphs in Fig. 2 (except for (e)). Our prime goal was not to minimize the number of parameters in the model, as there will inevitably be many variables governing performance of a high-level task even within a control group. Rather, we aimed to minimize the number of parameter-changes needed to account for the difference between ADHD and control children.

We omitted many aspects of MFFT performance, including frustration, inter-subject differences, search strategies, and trial-to-trial improvement (and true learning from temporal differences). We treated visual recognition and time simplistically. These simplifications certainly detract from the realism of the model. Explanations that are more complicated, however, are not needed until experimental results show the current model to be inadequate. The simplifications make it feasible to produce rigorous statements about whether simple aspects of reward processing or cognition are *sufficient* to explain the experimental results in ADHD.

Future work

The most obvious prediction from a model is that the structure of the model reflects the structure of reality.

From the interactions of cognitive and state-regulation mechanisms in the current model, we predict that performance in the Stop Signal Task and the Choice Delay Task (Solanto et al., 2001) will both individually and additively predict performance in the MFFT. Short-trial performance seems likely to be impaired in a very wide range of cognitive deficits, but hyperactivity caused by increased reward density (perhaps via Action Bias, as a state-regulation phenomenon), has been reported in several studies and such children may be a clinical group requiring distinct treatment.

We can also use the incompletely specified parts of the model to suggest areas for more open-ended exploration. For example, working memory deficits are very vague in the current ADHD literature: memories do not just blur but fade, transform, merge, disappear, or are superseded. Dopamine appetite needs to be looked for, not only by its explanatory power in computational models of behavioral data but also by studies of painfully bored animals and children.

The basic TD model was previously validated against the Choice-Delay Task (of Sonuga-Barke et al., 1992; see Williams & Dayan, 2004). The extended TD model has greater explanatory power than the basic model; the former has now been validated implicitly on the CDT, as well as explicitly on the MFFT. We expect that as the TD model of ADHD reaches maturity, it will become able to accommodate new paradigms without the need for extensions like those introduced in this paper. Then, rather than defining ADHD as a collection of symptoms or a constellation of rarer disorders, we will have a model of interactions among learning, appetites, and cognition, which is able to account for the multifarious symptoms of 'ADHD'.

ACKNOWLEDGMENTS

We are very grateful to Peter Dayan for his very helpful comments on an earlier draft of this paper.

REFERENCES

Alexander GE, DeLong MR, Strick PL. 1986; Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9: 357–381.

Aloise PA, Miller PH. 1991. Discounting in pre-schoolers: effect of type of reward agent. J Exp Child Psychol 52: 70–86.

American Psychiatric Association. 2000. Diagnostic and statistical manual of mental disorders: DSM-IV-TR. Washington DC, USA: APA.

Barkley RA, Anastopoulos AD, Guevremont DC, Fletcher KE. 1991. Adolescents with ADHD: patterns of behavioral adjustment, academic functioning, and treatment utilization. J Am Acad Child Adolesc Psychiatry 30: 752–761.

Barkley RA, Edwards G, Laneri M, Fletcher K, Metevia L. 2001. Executive functioning, temporal discounting, and sense of time in adolescents with attention deficit hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). J Abnorm Child Psychol 29: 541–556.

Barto AG, Sutton RS, Anderson CW. 1983. Neuron-like adaptive elements that can solve difficult learning control problems. IEEE Trans Systems, Man, and Cybernetics 13: 834–846.

Berman T, Douglas VI, Barr RG. 1999. Effects of methylphenidate on complex cognitive processing in attention-deficit hyperactivity disorder. J Abnorm Psychol 108: 90–105.

Blum K, Braverman ER, Holder JM, Lubar JF, Monastra VJ, Miller D, et al. 2000. Reward deficiency syndrome: a biogenetic model for the diagnosis and treatment of impulsive, addictive, and compulsive behaviors. J Psychoactive Drugs 32 Suppl i: 112.

Bush G, Frazier JA, Rauch SL, Seidman LJ, Whalen PJ, Jenike MA, et al. 1999. A nterior cingulate cortex dysfunction in attention-deficit/hyperactivity disorder revealed by fMRI and the Counting Stroop. Biol Psychiatry 45: 1542–1552.

Carlson CL, Tamm L. 2000. Responsiveness of children with attention deficit-hyperactivity disorder to reward and response cost: differential impact on performance and motivation. J Consult Clin Psychol 68: 73–83.

Cloninger CR. 1987. A systematic method for clinical description and classification of personality variants. A proposal. Arch Gen Psychiatry 44: 573–588.

Crean JP, de Wit H, Richards JB. 2000. Reward discounting as a measure of impulsive behavior in a psychiatric outpatient population. Exp Clin Psychopharmacol 8: 155–162.

Dalby JT, Kinsbourne M, Swanson JM, Sobol MP. 1977. Hyperactive children's underuse of learning time: correction by stimulant treatment. Child Dev 48: 1448–1453.

Douglas VI, Parry PA. 1994. Effects of reward and nonreward on frustration and attention in attention deficit disorder. J Abnorm Child Psychol 22: 281–302.

Downey KK, Pomerleau CS, Pomerleau OF. 1996. Personality differences related to smoking and adult attention deficit hyperactivity disorder. J Subst Abuse 8: 129–135.

Doyle AE, Biederman J, Seidman LJ, Weber W, Faraone SV. 2000. Diagnostic efficiency of neuropsychological test scores for discriminating boys with and without attention deficit-hyperactivity disorder. J Consult Clin Psychol 68: 477–488.

DuPaul G J, McGoey K E, Eckert T L, VanBrakle J. 2001. Preschool children with attention-deficit/hyperactivity disorder: impairments in behavioral, social, and school functioning. J Am Acad Child Adolesc Psychiatry 40: 508–515.

Elia J, Welsh PA, Gullotta CS, Rapoport JL. 1993. Classroom academic performance: improvement with both methylphenidate and dextroamphetamine in ADHD boys. J Child Psychol Psychiatry 34: 785–804.

Evans SW, Pelham WE, Smith BH, Bukstein O, Gnagy EM, Greiner AR, et al. 2001. Dose-response effects of methylphenidate on ecologically valid measures of academic performance and classroom behavior in adolescents with ADHD. Exp Clin Psychopharmacol 9: 163–175.

Faraone SV, Biederman J, Lehman B K, Spencer T, 1993. Norman D, Seidman LJ, et al. Intellectual performance and school failure in children with attention deficit hyperactivity disorder and in their siblings. J Abnorm Psychol 102: 616–623.

Felicetti DM, Julliard K. 2000. Behaviors of children with and without attention deficit hyperactivity disorder during a dental recall visit. ASDC J Dent Child 67: 246–9, 231.

Fried I, Wilson CL, Morrow JW, Cameron KA, Behnke ED, Ackerson LC, et al. 2001. Increased dopamine release in the human amygdala during performance of cognitive tasks. Nat Neurosci 4: 201–206.

Friston KJ, Tononi G, Reeke GN, Jr., Sporns O, Edelman GM. 1994. Value-dependent selection in the brain: simulation in a synthetic neural model. Neuroscience 59: 229–243.

Gilger JW, Kaplan BJ. 2001. Atypical brain develop-

ment: a conceptual framework for understanding developmental learning disabilities. Dev Neuropsychol 20: 465–481.

Gillberg C. 2003. Deficits in attention, motor control, and perception: a brief review. Arch Dis Child 88: 904–910.

Grodzinsky GM, Barkley RA. 1999. Predictive power of frontal lobe tests in the diagnosis of attention deficit hyperactivity disorder. Clin Neuropsychol 13: 12–21.

Haenlein M, Caul WF. 1987. Attention deficit disorder with hyperactivity: A specific hypothesis of reward dysfunction. J Am Acad Child Adolesc Psychiatry 26: 356–362.

Heiligenstein E, Guenther G, Levy A, Savino F, Fulwiler J. 1999. Psychological and academic functioning in college students with attention deficit hyperactivity disorder. J Am Coll Health 47: 181–185.

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

Hungund BL, Perel JM, Hurwic MJ, Sverd J, Winsberg BG. 1979. Pharmacokinetics of methylphenidate in hyperkinetic children. Br J Clin Pharmacol 8: 571–576.

Iaboni F, Douglas VI, Baker AG. 1995. Effects of reward and response costs on inhibition in ADHD children. J Abnorm Psychol 104: 232–240.

Jensen PS, Hinshaw SP, Swanson JM, Greenhill LL, Conners CK, Arnold LE, et al. 2001. Findings from the NIMH Multimodal Treatment Study of ADHD (MTA): implications and applications for primary care providers. J Dev Behav Pediatr 22: 60–73.

Karatekin C, Asarnow RF. 1998. Working memory in childhood-onset schizophrenia and attention-deficit/hyperactivity disorder. Psychiatry Res 80: 165–176.

Kempton S, Vance A, Maruff P, Luk E, Costin J, Pantelis C. 1999. Executive function and attention deficit hyperactivity disorder: stimulant medication and better executive function performance in children. Psychol Med 29: 527–538.

Kollins SH. 2003. Delay discounting is associated with substance use in college students. Addict Behav 28: 1167–1173.

Ljungberg T, Apicella P, Schultz W. 1992. Responses of monkey dopamine neurons during learning of behavioral reactions. J Neurophysiol 67: 145–163.

Lockwood KA, Marcotte AC, Stern C. 2001. Differentiation of attention-deficit/hyperactivity disorder subtypes: application of a neuropsychological model of attention. J Clin Exp Neuropsychol 23: 317–330.

Malhotra AK, Goldman D. 2000. The dopamine D(4) receptor gene and novelty seeking. Am J Psychiatry 157: 1885–1886.

Malone MA, Swanson JM. 1993. Effects of methylphenidate on impulsive responding in children with attention-deficit hyperactivity disorder. J Child Neurol 8: 157–163.

Mason DJ, Humphreys GW, Kent LS. 2003. Exploring selective attention in ADHD: visual search through space and time. J Child Psychol Psychiatry 44: 1158–1176.

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

O'Toole K, Abramowitz A, Morris R, Dulcan M. 1997. Effects of methylphenidate on attention and nonverbal learning in children with attention–deficit hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 36: 531–538.

Oie M, Rund BR. 1999. Neuropsychological deficits in adolescent-onset schizophrenia compared with attention deficit hyperactivity disorder. Am J Psychiatry 156: 1216–1222.

Raggio DJ. 1999. Visuomotor perception in children with attention deficit hyperactivity disorder—combined type. Percept Mot Skills 88: 448–450.

Rolls ET, Treves A. 1998. Neural Networks and Brain Function. Oxford, UK: Oxford University Press.

Rubia K, Overmeyer S, Taylor E, Brammer M, Williams SC, Simmons A, et al. 1999. Hypofrontality in attention deficit hyperactivity disorder during higher-order motor control: a study with functional MRI. Am J Psychiatry 156: 891–896.

Sagvolden T, Aase H, Johansen E. 2000. The neuropsychology of attention-deficit/hyperactivity disorder (AD/HD) in a hypodopaminergic system. 6th Internet World Congress for Biomedical Sciences, INABIS.

Sagvolden T, Aase H, Zeiner P, Berger D. 1998. Altered reinforcement mechanisms in attention–deficit/hyperactivity disorder. Behav Brain Res 94: 61–71.

Scheres A, Oosterlaan J, Sergeant JA. 2001. Response inhibition in children with DSM-IV subtypes of AD/HD and related disruptive

disorders: the role of reward. Neuropsychol Dev Cogn Sect C Child Neuropsychol 7: 172–189.

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

Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. Science 275: 1593–1599.

Shea SA. 1996. Behavioural and arousal-related influences on breathing in humans. Exp Physiol 81: 1–26.

Sleator EK, Ullmann RK. 1981. Can the physician diagnose hyperactivity in the office? Pediatrics 67: 13–17.

Solanto MV, Abikoff H, Sonuga-Barke E, Schachar R, Logan GD, Wigal T, et al. 2001. The ecological validity of delay aversion and response inhibition as measures of impulsivity in AD/HD: a supplement to the NIMH multimodal treatment study of AD/HD. J Abnorm Child Psychol 29: 215–228.

Sonuga-Barke EJ. 2002a. Interval length and time–use by children with AD/HD: a comparison of four models. J Abnorm Child Psychol 30: 257–264.

Sonuga–Barke E J, Dalen L, Remington B. 2003. Do executive deficits and delay aversion make independent contributions to preschool attention–deficit/hyperactivity disorder symptoms? J Am Acad Child Adolesc Psychiatry 42: 1335–1342.

Sonuga-Barke EJ, Taylor E, Sembi S, Smith J. 1992. Hyperactivity and delay aversion-I. The effect of delay on choice. J Child Psychol Psychiatry 33: 387–398.

Sonuga-Barke EJS. 2002b. Psychological heterogeneity in AD/HD—a dual pathway model of behavior and cognition. Behav Brain Res 130: 29–36.

Sunohara G A, Malone M A, Rovet J, Humphries T, Roberts W, Taylor MJ. 1999. Effect of methylphenidate on attention in children with attention deficit hyperactivity disorder (ADHD): ERP evidence. Neuropsychopharmacology 21: 218–228.

Swanson JM, Flodman P, Kennedy J, Spence MA, Moyzis R, Schuck S, et al. 2000a. Dopamine genes and ADHD. Neurosci Biobehav Rev 24: 21–25.

Swanson J, Oosterlaan J, Murias M, Schuck S, Flodman P, Spence MA. et al. 2000b. 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 USA 97: 4754–4759.

Tannock R, Schachar R, Logan G. 1995. Methylphenidate and cognitive flexibility: dissociated dose effects in hyperactive children. J Abnorm Child Psychol 23: 235–266.

Tannock R, Schachar RJ, Carr RP, Logan GD. 1989. Dose-response effects of methylphenidate on academic performance and overt behavior in hyperactive children. Pediatrics 84: 648–657.

van der Meere J, Gunning B, Stemerdink N. 1999. The effect of methylphenidate and clonidine on response inhibition and state regulation in children with ADHD. J Child Psychol Psychiatry 40: 291–298.

Volkow ND, Wang GJ, Fowler JS, Gatley SJ, Logan J, Ding YS, et al. 1998. Dopamine transporter occupancies in the human brain induced by therapeutic doses of oral methylphenidate. Am J Psychiatry 155: 1325–1331.

Watkins C. Learning from Delayed Rewards. 1989. University of Cambridge.

Wilson M, Daly M. 2003. Do pretty women inspire men to discount the future? Proc R Soc Lond B doi:10.1098/rsbl.2003.0134 (published online).

Young SE, Stallings MC, Corley RP, Krauter KS, Hewitt JK. 2000. Genetic and environmental influences on behavioral disinhibition. Am J Med Genet 96: 684–695.

APPENDIX

THE TEMPORAL DIFFERENCE (TD) MODEL

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

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

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

Montague et al. (1996) used exactly $V(t)$ as the critic in a form of the actor-critic architecture (Barto et al., 1983). Here, we use $V(t) = Q(t,u)$, where $Q(t,u)$ depends (Watkins, 1989) on the action u chosen by the model at time t, where u is either a or w, for acting or waiting. In our parameterization,

$$s_u(t) = Q(t,u).$$

The error-signal is used to learn $Q(t,u)$, and these values are used to control action selection.

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 $s_u(t)$ is

$$s_u \longrightarrow s_u + \eta\delta(t)$$

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

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 his behavior to encourage exploration of explore various behavior patterns, by using the softmax function

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

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.

ALGORITHM IMPLEMENTING THE MFFT TASK FOR MODEL B

The model includes a vector c, indicating the current best estimate of the probability that each of the target pictures matches the index picture. Difficulty levels of the trials are selected from a beta distribution with parameters 0.5 and 1, to restrict the number of difficult trials.

The simulated dopamine level is incremented by 1 at the end of each trial if the model has produced the right answer. The level then decays exponentially, being multiplied at each time-step by 0.9. For the two-tailed case, with Dopamine set-point DaSET (0.7), at each time-step in which the current Dopamine level Da is outside the range DaSET ± 0.3, the effective Action Bias ϕ_e is calculated as:

$$\phi_e = \phi + \theta\,|\,\mathrm{DaSET} - \mathrm{Da}\,|$$

where ϕ is the standard action bias for this experimental group, and θ is the gearing between dopamine and Action Bias [shown as "two-tail AB" in Fig. 2(c)]. Whenever the maximum array entry exceeds the certainty threshold (0.99), the

model ceases all inspection of pictures in for the remainder of the trial. During each time-step the model spends away from the task, c is degraded according to Δ (default 0.008) as shown below:

```
Da←0
Repeat (for each of 5000 simulated trials, all the same length)
    Select trial difficulty d from beta(.5,1)
    previouspicturenumber←nul
    c←number of pictures^-1
    p_t←the number of the target picture
    thinkdone←false
    Repeat (for each timestep t in the trial)
        if thinkdone=true, choose to play; otherwise:
          Using softmax, select action from {work systematically; work unsystematically; play},
          based on heuristic expected rewards {reward discounted from end of trial; half that; φ_e }
        if working:
                choose a picture p to inspect:
                  (either systematically: previouspicturenumber+1; or unsystematically: random)
                if rand*d<0.25 (i.e. the comparison of difficulty d produces some information)
                    if p= p_t: c_{i=p} ←1; c_{i≠p}←0
                      else: c_{i=p} ←0
        c ← c / Σ( c)
        if max(c)>0.99 thinkdone←true
        c ← (c+Δ) / Σ( c+Δ)
        Da←Da * 0.9
        previouspicturenumber←p
    if max(c) corresponds to index picture: Da←Da+1
```