A formal investigation of dopamine's role in Attention-Deficit/Hyperactive Disorder: Evidence for asymmetrically effective reinforcement learning signals

by

Jeffrey Cockburn
B.Sc, University of Victoria, 2005

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in Interdisciplinary Studies (Computer Science, Psychology)

Supervisory Committee

A formal investigation of dopamine's role in Attention-Deficit/Hyperactive Disorder: Evidence for asymmetrically effective reinforcement learning signals

by

Jeffrey Cockburn
B.Sc, University of Victoria, 2005

Supervisory Committee

Dr. Clay Holroyd (Department of Psychology)
Co-Supervisor

Dr. Jens Weber (Department of Computer Science)
Co-Supervisor

Dr. Tony Marley (Department of Psychology)
Departmental Member

Abstract

Supervisory Committee
Dr. Clay Holroyd (Department of Psychology)
Co-Supervisor
Dr. Jens Weber (Department of Computer Science)
Co-Supervisor
Dr. Tony Marley (Department of Psychology)
Departmental Member

Attention-Deficit/Hyperactive Disorder is a well studied but poorly understood disorder. Given that the underlying neurological mechanisms involved in the disorder have yet to be established, diagnosis is dependent upon behavioural markers. However, recent research has begun to associate a dopamine system dysfunction with ADHD; though, consensus on the nature of dopamine's role in ADHD has yet to be established. Here, I use a computational modelling approach to investigate two opposing theories of the dopaminergic dysfunction in ADHD. The hyper-active dopamine theory posits that ADHD is associated with a midbrain dopamine system that produces abnormally large prediction errors signals; whereas the dynamic developmental theory argues that abnormally small prediction errors give rise to ADHD. Given that these two theories center on the size of prediction errors encoded by the midbrain dopamine system, I have formally investigated the implications of each theory within the framework of temporal-difference learning, a reinforcement learning algorithm demonstrated to model midbrain dopamine activity. The results presented in this thesis suggest that neither theory provides a good account for the behaviour of children and animal models of ADHD. Instead, my results suggest ADHD is the result of asymmetrically effective reinforcement learning signals encoded by the midbrain dopamine system. More specifically, the model presented here reproduced behaviours associated with ADHD when positive prediction errors were more effective than negative prediction errors. The biological sources of this asymmetry are considered, as are other computational models of ADHD.

Table of Contents

List of Tables

List of Figures

Acknowledgments

First and foremost, I would like to thank Dr. Clay Holroyd. Clay has contributed in so many ways to this research and to my development as a researcher it simply cannot be adequately acknowledged here. In short, over the past two years Clay has patiently taught me more about science and how to think about the world than I thought possible. Dr Jens Weber, who has encouraged my meandering research endeavours since I was an undergraduate, has provided me with a consistently clear-minded approach to research from which to root myself. I would also like to thank Dr. Jim Tanaka for supporting me and providing every opportunity to develop as a scientist.

I can't thank my family enough for encouraging me in my seemingly endless scholastic career. Your nurturing support has provided an unshakable foundation from which to explore the world. Finally, words cannot express my thanks to Paulina, my fiancée; for reminding me that love is not just a parameter.

Dedication

To my Grandparents: Grandpa and Granny (Robert and Edna Cockburn), Granddad and Grandma (Charlie and Betty Pierce).

1 Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is the most common childhood onset disorder encountered in primary care settings (Sutker & Adams, 2001). Approximately 3-7% of the population fall within diagnostic criteria, which include developmentally inappropriate hyperactive, impulsive and inattentive behaviour (American Psychiatric Association, 2000). Though these behavioural symptoms are poorly operationalized, they are commonly observed as difficulty staying seated, excessive motor movement, difficulty waiting one's turn, and excessive manipulation of objects (Barkley, 1997a). ADHD typically manifests itself before the age of 7, is persistent, and is present in two or more settings (e.g. home and school) (American Psychiatric Association, 2000). Longitudinal research has shown that approximately 50-80% of children with ADHD will continue to have the disorder into adolescence, with 30-50% continuing on into adulthood (Barkley, 1997b). In severe cases, social and psychological development are at risk (Scahill & Schwab-Stone, 2000; Taylor, 1994) which likely contributes to a higher incidence of social dysfunction and substance abuse as adults (Wilens, Biederman, & Spencer, 2002; Wilens, Faraone, & Biederman, 2004).

Many of the challenging behaviours associated with ADHD resemble those seen in individuals with frontal lobe pathology (Barkley, 1997a; Chelune, Ferguson, Koon, & Dickey, 1986), which emphasize deficits in executive control (J. A. Sergeant, 2005; Sonuga-Barke, 2002b). Neuropsychological investigations have implicated the frontal lobe in ADHD (Gorenstein, Mammato, & Sandy, 1989; Grodzinsky, & Diamond, 1992), while both functional and structural neuroimaging studies point toward a dysfunction within frontal-striatal circuits (Casey et al., 1997; Lou, Henriksen, Bruhn, Borner, & Nielsen, 1989). Reduced frontal activation has been observed in children with ADHD during Stroop, stop and motor priming tasks (Bush et al., 1999; Rubia et al., 1999). Children with ADHD also make smaller adjustments following errors than control subjects during speeded response time tasks (Schachar et al., 2004; J. A. Sergeant & van

der Meere, 1988), while administration of methylphenidate has been found to normalize post-error adjustments (Solanto, 1990).

Despite decades of research, a causal neurological model of the ADHD has yet to be established (Coghill, Nigg, Rothenberger, Sonuga-Barke, & Tannock, 2005); hence, a behavioural description of the disorder is still necessary. The deficits associated with ADHD are generally placed under the rubric of a dysfunctional executive control system. As helpful as this may be in guiding diagnosis, treatment and our understanding of the disorder, it also over-generalizes the disorder, potentially clouding a more pointed investigation. However, recent research on ADHD has begun to frame the disorder as a dysfunction rooted in the midbrain dopamine system. Suggestion of dopaminergic dysfunction originally stem from the paradoxical finding that stimulants, such as methylphenidate, help to alleviate the symptoms associated with ADHD (Bradley, 1937; Vitiello et al., 2001). By increasing extracellular dopamine levels, which act to *induce* ADHD-like symptoms in normal subjects, dopamine agonists normalizes the hyperactive, impulsive and inattentive behaviours commonly associated with ADHD (Castellanos, 1997; Castellanos & Tannock, 2002; Grace, 2001). Following this line of research, recent genetic studies have associated ADHD with polymorphic sites at several dopamine-related genes involving dopamine receptors (Swanson et al., 2000), degradation (Bellgrove et al., 2005), and transport (Krause, Dresel, Krause, la Fougere, & Ackenheil, 2003). Finally, decades of behavioural studies have demonstrated atypical reinforcement learning in ADHD (Luman, Oosterlaan, & Sergeant, 2005). It should be noted that much of the behavioural evidence provides complex and sometimes inconsistent results, but together with a better understanding of the midbrain dopamine system's role in reinforcement learning behavioural studies further support the dopaminergic dysfunction theory of ADHD.

1.1 The midbrain dopamine system

Dopamine's role is not adequately described in terms of excitation or inhibition as with most neurotransmitters; rather, its influence is captured more aptly as a neuromodulator capable of gating information by modulating the target neuron's gain or activation function (Grace, 2001; Servan-Schreiber, Printz, & Cohen, 1990). Hence, its effect is largely dependent on the target system and its current state (W. Schultz, 2002). Dopamine's impact on target neurons can be long lasting (Missale, Nash, Robinson, Jaber, & Caron, 1998), extending the temporal window during which coincidence detection can occur (Gray, Feldon, Rawlins, Hemsley, & Smith, 1991). This is thought to play a crucial role in both long term potentiation (LTP) (Pedarzani & Storm, 1995) and long term depression (LTD) (Sajikumar & Frey, 2004). This is facilitated in frontal cortex (Goldman-Rakic, Leranth, Williams, Mons, & Geffard, 1989) and the basal ganglia (Smith & Bolam, 1990) via "synaptic triads" in which dopamine terminals make contact with a synapse. This acts to implement a 3-factor learning rule in which synaptic connections are strengthened only if the pre-synaptic, post-synaptic and dopamine neurons are simultaneously activated. Thus, dopamine function is crucial for learning and action selection.

Dopamine neurons in the midbrain dopamine system exhibit two distinct modes of activity: tonic and phasic. Tonic activity is the base level firing rate of dopamine neurons. Characterized by a moderate and consistent activation level, tonic activity is largely regulated by extracellular dopamine concentrations and glutamate release from frontal afferents in close proximity to dopamine terminals (Grace, 2001). However, afferent drive can push the midbrain dopamine system into a phasic mode of activity (Pucak & Grace, 1994). Phasic activity is characterized by a transient increase or decrease in neuronal spike-rate producing a rapid burst or dip in dopamine release into the synaptic cleft (W. Schultz, Dayan, & Montague, 1997). Phasic bursts are observed after unexpected rewards while phasic dips occur when an expected reward is withheld (W. Schultz et al., 1997). As learning occurs, phasic activity shifts from the time of reward delivery to the time of a stimulus predicting future reward (W. Schultz et al., 1997).

These transient increases and decreases in dopamine have been hypothesized to encode an error in the prediction of a reward (Montague, Dayan, & Sejnowski, 1996; W. Schultz et al., 1997). A considerable body of data has been shown to be compatible with a temporal-difference reinforcement learning model of dopamine in which the midbrain dopamine system encodes a reward prediction error that can be used by target systems for reinforcement learning (Montague et al., 1996; A. D. Redish, 2004; W. Schultz et al., 1997; W. Schultz, 1998; Waelti, Dickinson, & Schultz, 2001).

1.2 Temporal-difference reinforcement learning

Reinforcement learning provides a computational framework through which we can formally analyze how animals learn to predict the outcome of actions and events, and how they behave so as to optimize rewards. My proposed model of ADHD is based on a temporal-difference reinforcement learning (TDRL) algorithm in which actions are selected so as to maximize the discounted value of future rewards (Sutton & Barto, 1998). Actions are selected based on a value function *V(s)*, which is defined as the expected future rewards discounted by the delay between the current state, s_t, and the time of reward:

$$V(s) = E\left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \middle| s_t = s \right\} \tag{1.1}$$

where $E\{\ \}$ denotes the expected value, t is the current time-step, r is the expected reward $k+1$ time-steps in the future, and γ is a discounting factor ($0 \leq \gamma \leq 1$). The discount factor, γ, determines the degree to which reward delay is considered in the state value. Given the choice between two mutually exclusive rewards, r_1 and r_2, where r_2 is delivered later than r_1, the discount factor can be used to help decide which reward should be pursued. Small discount factors will impose a large penalty for reward delay; hence, immediate rewards will be preferentially selected over delayed rewards even if the delayed reward is larger.

As the discount factor increases, $\gamma \rightarrow 1$, the penalty for reward delay is reduced, and as a result, the state value will increasingly include the value of future rewards. Hence, one can think of the discount factor as determining how near- or far-sighted state values are.

In TDRL, the world is represented as a finite set of discrete states referred to as the state-space. A learning agent experiences the world by transitioning from state to state, either because of an action taken by the agent or due to some processes in the world outside the agent's control which forces it into a new state. The goal of TDRL is to learn the value of each state in the state-space. This is accomplished by computing a numerical discrepancy between the expected outcome of a state and the actual outcome whenever a state transition is made, referred to as a reward prediction error. This error signal can be used to minimize the difference between predicted and observed outcomes such that the agent can accurately predict the outcome of being in a given state, and hence, can act in a way that maximizes the amount of reward received over time (Sutton & Barto, 1998).

The value function, $V(s)$, is learned by calculating two equations at each state transition. When the agent leaves state s_t and enters state s_{t+1}, at which time it receives a reward r_{t+1} from the environment, the reward prediction error is defined as:

$$\delta_t = r_{t+1} + \gamma V(s_{t+1}) - V(s_t) \tag{1.2}$$

where γ is the discount factor ($0 \leq \gamma \leq 1$) determining how near- or far-sighted the prediction error calculation should be. The value function is then updated for state s_t by:

$$V(s_t) = V(s_t) + \eta \delta_t \tag{1.3}$$

where η is a learning rate parameter ($0 \leq \eta \leq 1$) specifying the proportion of the δ_t signal that should be absorbed by the state value.

A mechanism commonly used to improve TDRL efficiency is an eligibility trace (Sutton & Barto, 1998). Rather than update the value of a single state after each transition, an eligibility trace provides a form of state transition memory such that the set of all previously encountered states have their values updated after each transition. A replacing eligibility trace is defined as:

$$e(s_t) = \begin{cases} \gamma\lambda e(s_{t-1}) & \text{if } s \neq s_t \\ 1 & \text{if } s = s_t \end{cases} \tag{1.4}$$

where γ is the discount factor previously discussed, ($0 \leq \gamma \leq 1$), and λ is the trace decay rate, ($0 \leq \lambda \leq 1$). A state's "memory" is turned on whenever it is encountered by setting its trace weight to 1. Once on, the state's "memory" wanes as a function of transitions following that state, decaying by a factor of $\gamma\lambda$[1] after each state transition. In this way, recent decisions assume more of the praise (or blame) when rewards (or punishments) are encountered and state values are updated accordingly. In order to employ an eligibility trace, a slightly modified value update equation is required, defined as:

$$V(s_t) = V(s_t) + \eta\delta_t e(s_t) \tag{1.5}$$

where the set of all previously encountered states have their values updated according to the trace strength defined by $e(s_t)$.

Learning is driven by the reward prediction error signal, δ. If the outcome of a state transition turns out better than expected, then $\delta_t > 0$ and the value of the initial state is increased so that its value more accurately predicts the positive outcome. If, on the other

[1] Trace decay is a factor of both γ and λ parameters. λ defines the "memory" decay rate after each transition, while γ accounts for the discounted value of rewards encountered at future states.

hand, the observed value is worse than expected, then $\delta_t < 0$ and the initial state value is decreased to compensate for its overly optimistic prediction. Finally, if the value of the initial state completely predicts the observed outcome, then $\delta_t = 0$, indicating that the situation is well learned and no changes need be made.

Of particular importance, note that the δ signal transfers values backwards from rewarding states to anticipatory states, chaining state values together such that future rewards can be predicted. By chaining state values together, TDRL is capable of crediting actions and the temporal relationships among them with the outcome they produce. Hence, TDRL provides a solution to the temporal credit assignment problem; that is, how one should distribute credit among the many decisions that lead to success or failure. Earlier models such as the Rescorla-Wagner learning rule, which has been widely used to model conditioning behaviour, assigned credit equally among all decisions (Miller, Barnet, & Grahame, 1995). While this may be a fair assumption for simple conditioning paradigms, it fails when more complex action sequences are involved or temporal discrimination is required.

Phasic activity in the midbrain dopamine system exhibits behaviour remarkably similar to that of a TDRL reward prediction error and is thought to facilitate a learning process in much the same way (Montague et al., 1996; W. Schultz et al., 1997). Phasic dopamine bursts occur when events turn out better than expected, encoding a positive reward prediction error, while phasic dopamine dips that occur when events are worse than expected encode negative reward prediction errors. Targets of the midbrain dopamine system use these error signals to improve task performance by strengthening neural activity that lead to rewards and weakening activity that did not.

1.3 Dopamine's role in ADHD

As was previously discussed, drug, genetic and behavioural studies suggest an association between ADHD and dopamine system dysfunction. Stimulants such as

methylphenidate were found to effectively reduce the problematic behaviour associated with ADHD (Bradley, 1937; Vitiello et al., 2001). These drugs are known to act by increasing extracellular concentrations of dopamine (Cooper, Bloom, & Roth, 2003). This finding was particularly perplexing. Typically, stimulants *induce* precisely the hyperactive, impulsive and inattentive behaviours associated with ADHD. So why do stimulants reduce precisely the symptoms they act to induce if those symptoms are present prior to drug treatment?

1.4 Hyper-active dopamine theory of ADHD

Stemming from a pathophysiological investigation of this paradoxical finding, A. Grace (2001) proposed the hyper-active dopamine theory of ADHD. Based on an investigation of dopamine agonists typically used to treat ADHD, which act to increase extra-cellular concentrations of dopamine, the proposal holds that ADHD must be associated with low concentrations of extra-cellular dopamine in striatum. This, it is argued, is likely related to reduced stimulation from frontal afferents, which have been shown to modulate striatal concentrations of extra-cellular dopamine (Floresco, West, Ash, Moore, & Grace, 2003; Grace, 1991). Extra-cellular concentrations of dopamine have been shown to modulate phasic activity in the midbrain dopamine system (Grace, 1991). This modulation takes place via inhibitory dopamine autoreceptors in the extra-synaptic space at dopamine neuron terminals (Nowycky & Roth, 1977) and cell bodies (Pucak & Grace, 1994). As extracellular concentrations of dopamine increase, dopamine neurons are increasingly inhibited, while decreased concentrations lead to decreased inhibition (Figure 1.1). Hence, extra-cellular dopamine acts to down-regulate phasic activity in the midbrain dopamine system.

Decreased inhibition via low concentrations of extra-cellular dopamine will push the dopamine system into a hyper-active state characterized by exaggerated phasic activity (Grace, 2001). Given that phasic activity is thought to convey a neural reinforcement learning signal (W. Schultz et al., 1997), abnormal phasic activity will lead to

reinforcement learning deficits, which have been demonstrated in children with ADHD (Luman et al., 2005). Dopamine agonists used to treat ADHD, such as methylphenidate, act by increasing extracellular concentrations of dopamine via transporter blockade (Cooper et al., 2003), which normalizes inhibition on the midbrain dopamine system (Grace, 2001). This, in turn, will pull the midbrain dopamine system's phasic activity back into the typical range, normalizing the reinforcement learning signal. In summary, the hyper-active dopamine theory of ADHD holds that ADHD is the result of abnormally large prediction errors, which occur because low concentrations of extra-cellular dopamine are unable to properly down-regulate the midbrain dopamine system (Grace, 2001).

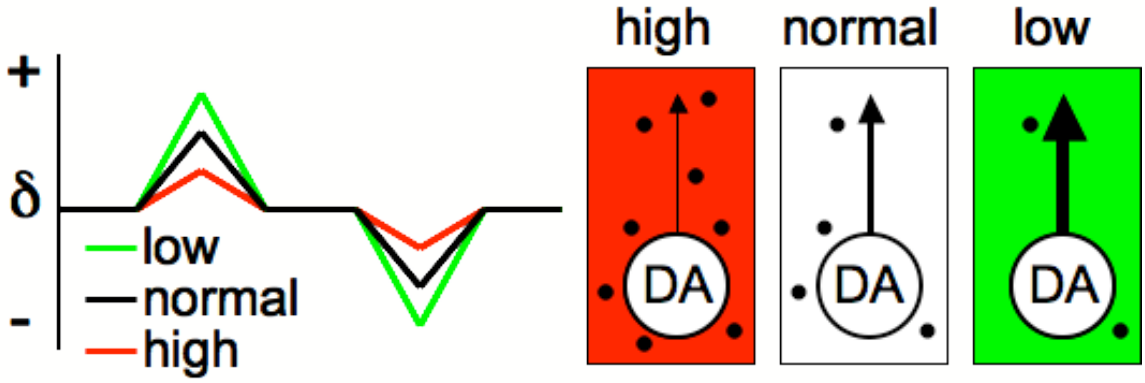

Figure 1.1: *Extracellular dopamine's impact on phasic dopamine activity* (Grace, 2001). High concentrations of extracellular dopamine (red box) inhibit dopamine neurons excessively, stunting phasic prediction error signals coming from the midbrain dopamine system (red line). Low concentrations of extracellular dopamine (green box) disinhibit dopamine neurons, resulting in exaggerated prediction error signals (green line).

1.5 Hypo-active dopamine theory of ADHD

While the hyper-active dopamine theory of ADHD appears promising, T. Sagvolden et al. (2005) introduced the dynamic developmental theory of ADHD. Founded on behavioural investigations of the symptoms associated with ADHD in both an animal model of ADHD and diagnosed children (Sagvolden, Hendley, & Knardahl, 1992;

Sagvolden, Pettersen, & Larsen, 1993; Sagvolden, Aase, Zeiner, & Berger, 1998; Sagvolden, 2000; Sagvolden, Russell, Aase, Johansen, & Farshbaf, 2005), they conclude that ADHD is rooted in a dopamine dysfunction that results in stunted phasic activity in the midbrain dopamine system.

The relationship between the time interval separating behaviour and reinforcer, and the effect a reinforcer can have on behaviour is referred to as the delay-of-reinforcement gradient (Figure 1.2) (Catania, Sagvolden, & Keller, 1988). The reinforcing effect is largest for behaviours immediately preceding a reinforcer and wanes as a function of the temporal delay separating behaviour and reinforcer. The dynamic developmental theory of ADHD proposes that ADHD is associated with shorter and steeper delay-of-reinforcement, and therefore, inefficient reinforcement and extinction processes (Figure 1.2) (Sagvolden et al., 2005). This, it is argued, is likely due to a hypo-active dopamine system producing abnormally small prediction error signals.

While this proposal agrees with the hyper-active dopamine theory that ADHD is associated with low concentrations of extracellular dopamine, they argue that the normal coupling between tonic and phasic activity is disrupted such that children with ADHD have abnormally low tonic and phasic activation. A dysfunctional dopamine system will result in behaviours being learned slowly and less efficiently (W. Schultz, 2002), which is argued to produce the hyperactive, impulsive and inattentive behaviour associated with ADHD (Sagvolden et al., 2005).

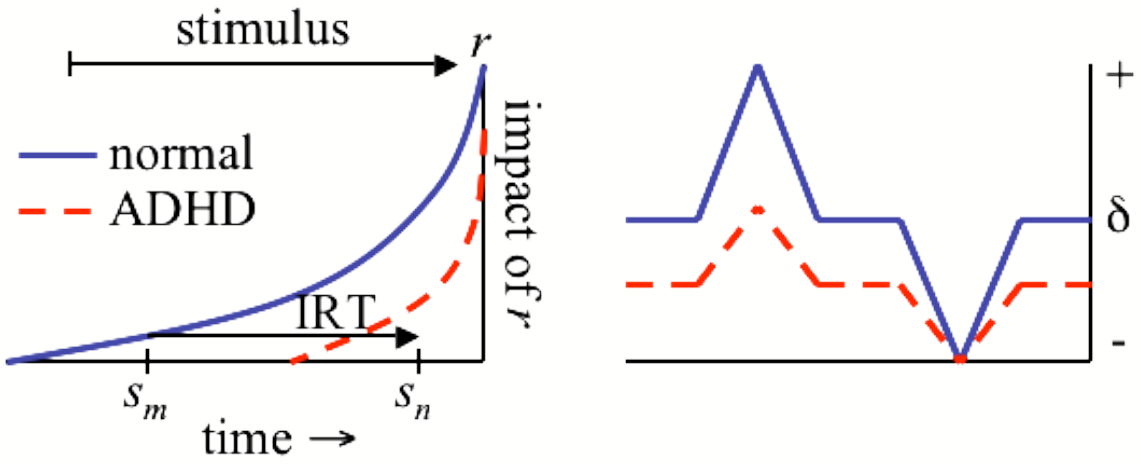

Figure 1.2: *The effect of δ's size on the delay-of-reinforcement gradient* (Sagvolden, Johansen, Aase, & Russell, 2005). The impact of a reinforcer, *r*, is largest for recent states, and wanes as a function of time (left). Thus, in a normal system state s_n will be impacted by *r* to a greater degree than state s_m. It is argued that ADHD is associated with stunted δ signals (right: dashed red line), resulting in a shorter and steeper delay-of-reinforcement gradient (left: dashed red line). A normal gradient allows reinforcer *r* to influence state s_m, though to a lesser degree than state s_n. A shorter and steeper gradient reduces the impact of *r* on s_n, while *r* is not able to influence s_m at all. This will result in slower and less efficient learning. A sufficiently long gradient is required for tertiary binding of stimulus, behaviour and reinforcer in order to support sustained attention (left: stimulus arrow). Finally, the delay-of-reinforcement gradient will determine the temporal range of IRTs that can be effectively sustained. (left: IRT arrow).

1.6 Computational Modelling of Neuromodulation in ADHD

Both the hyper- and hypo-active dopamine theories of ADHD hinge upon, and make directly opposing claims regarding phasic activity of the midbrain dopamine system. I investigate this conflict using TDRL as a normative computational framework in which phasic dopamine activity is modelled as a reward prediction error. Reinforcement learning provides a sound mathematical basis (Bertsekas & Tsitsiklis, 1996) with a close relationship to behavioural learning (Sutton & Barto, 1981). Furthermore, TDRL has been used extensively to simulate dopamine activity and behaviour (Dayan & Niv, 2008;

Montague et al., 1996; Niv, 2009; W. Schultz et al., 1997); hence it provides an ideal basis from which to base my investigation of dopaminergic dysfunction in ADHD.

Given this theoretical framework, the hyper- and hypo-active dopamine theories of ADHD hold that TDRL prediction errors, (δ in equation (1.2)), carried by the dopamine system are abnormally large and abnormally small respectively. The midbrain dopamine system targets multiple neural systems, each likely using δ in a unique way and each system influencing behaviour differently (C. B. Holroyd & Coles, 2002). Though the interaction among these systems will likely play a role in a detailed understanding of ADHD, the hyper-/hypo-active dopamine theories of ADHD limit their scope to activity of the midbrain dopamine system, largely ignoring efferent systems. Hence, to avoid over-complicating my proposed model, I collapse all dopaminergic targets into a single abstract structure that learns about its environment and drives behaviour.

I begin by reviewing two behavioural experiments that form the basis of the hypo-active dopamine theory of ADHD, followed by an outline of my proposed model and simulations. Next, I argue on computational grounds that neither abnormally large (hyper-active) nor abnormally small (hypo-active) reward prediction errors can account for reported ADHD behaviour. I suggest that these behaviours are best accounted for by an asymmetry between positive and negative reward prediction errors. Finally, the biological implications of this are examined and I compare my proposed model to other models of ADHD.

2 Methods

I aim to clarify the neurological mechanisms involved in ADHD by investigating two opposing theories that associate ADHD with dopamine dysfunction. The hyper-active dopamine theory proposes an association between ADHD and exaggerated phasic dopamine activity (Grace, 2001), whereas the hypo-active theory proposes the exact opposite; namely, an association between ADHD and stunted phasic dopamine activity (Sagvolden et al., 2005). I formally test these two theories by applying a TDRL model of the midbrain dopamine system to simulations of a multiple fixed-interval/extinction (multi-FI/EXT) schedule of reinforcement task. Activity of the midbrain dopamine system is modelled as a reward prediction error, which can be parametrically scaled to formally investigate the behavioural implications of abnormally large and small reward prediction errors. The model's behaviour is compared to reported behavioural results on multi-FI/EXT tasks for an animal model of ADHD, (Sagvolden et al., 1992) as well as diagnosed children (Sagvolden et al., 1998). In the following, I outline those behavioural tasks and results, then provide a detailed review of the structure and dynamics of my proposed model.

2.1 Behavioural experiments and results

The hypo-active dopamine theory of ADHD is based largely on performance measures of the multi-FI/EXT task outlined in Figure 2.1. A schedule is termed multiple when two or more components alternate, each in the presence of a different stimulus. The task is such that a reinforcer is delivered for the first response made after a fixed interval of time during the fixed-interval component. All responses made prior to this fixed delay deliver nothing. In a typical design, the fixed-interval component begins with a stimulus change, such as a light turning on and remaining on for the duration of the component. Each fixed-interval component consists of several trials, with each trial being terminated by the delivery of a reinforcer. Following a fixed-interval component's last trial, an extinction component begins with a corresponding stimulus change, such as a light turning off. No

reinforcers are delivered during an extinction component; hence, each extinction component consists of only a single trial. When an extinction component is complete the next fixed-interval component begins, signalled by its corresponding stimulus change. The fixed-interval component is argued to measure reactivity to reinforcers, while the extinction component measures sustained attention since the subject must use context to maintain sustained control of their behaviour. As such, a multi-FI/EXT task provides measures for: a) Hyperactivity, quantified as response rate during fixed-interval components, b) Motor impulsivity, quantified as inter-response times (IRTs) during fixed-interval components, and c) Sustained attention, quantified as response rate during extinction components (Sagvolden, 2000)[2].

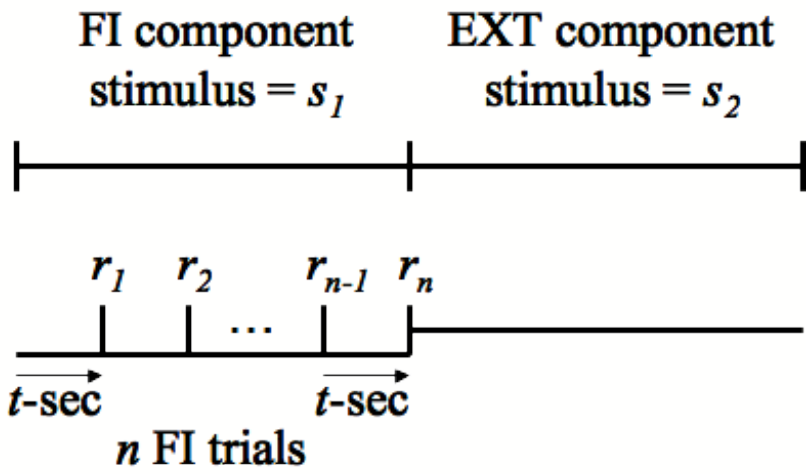

Figure 2.1: *General outline of the multi-FI/EXT task*. Fixed-interval (FI) and extinction (EXT) components alternate, each signalled by its own stimulus (e.g. light on/off). Each FI component consists of *n* trials. Each trial is terminated with a reinforcer, *r*, being delivered for the first response after a fixed delay time, *t*. No reinforcers are delivered during the EXT component.

Groups of control and children with ADHD naïve to medication as well as an animal model of ADHD have been tested on comparably similar multi-FI/EXT tasks (see

[2] These definitions of hyperactivity, impulsivity and attention are controversial (Alsop, 2007). I do not hold that they are the correct terms with which ADHD behaviour should be defined. I use these definitions only to be consistent with the literature.

(Sagvolden, 2000; Sagvolden et al., 2005) for a discussion relating animal and human studies). In the following I provide an overview of the methods and results for both animal and human experiments.

2.1.1 Animal experiment: methods and results

There are numerous studies investigating animal models of ADHD (Berger & Sagvolden, 1998; Boix, Qiao, Kolpus, & Sagvolden, 1998; Sagvolden et al., 1993; Wultz & Sagvolden, 1992) showing the spontaneously hypertensive rat (SHR) to provide the best model of the disorder (Sagvolden, 2000; Sagvolden et al., 2005). I focus on the experiment outlined in Sagvolden et al. (1992) due to the simplicity of the experimental design and because it provides a measure of stable response behaviour. Response behaviour of the control group, consisting of Wistar-Kyoto (WKY) rats, and an SHR ADHD-model group are compared. As illustrated in Figure 2.2, water deprived rats were subject to 58 sessions of a multi-FI/EXT task in which water acted as a reinforcer. Each session lasted approximately 1 hour, after which the animal was given free access to water for one hour, followed by 22 hours of water deprivation before the next session. The testing chamber light was turned on to signal a 2-minute fixed-interval component, during which the first lever press after 2-minutes was reinforced by a drop of water. The chamber light was turned off to signal a 5-minute extinction component, during which no water was delivered. Each session consisted of four consecutive components with no breaks: 1) A 2-minute fixed-interval component in which a maximum of 7 reinforcers were delivered, 2) a 5-minute extinction component, 3) a second 2-minute fixed-interval component identical to the first, and 4) a final 5-minute extinction component to complete the session.

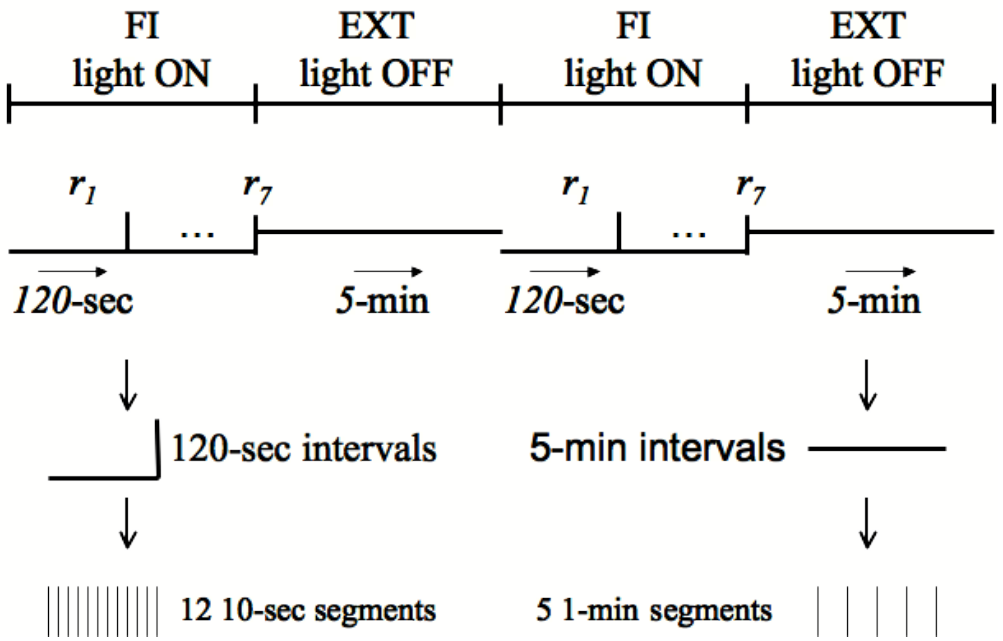

Figure 2.2: *Multi-FI/EXT animal task* (Sagvolden et al., 1992). Each of the 58 sessions consist of 4-components: FI→EXT→FI→EXT. In each FI component a maximum of seven reinforcers are delivered for the first response after a fixed delay of 120 seconds. The 120-second FI intervals are divided into twelve 10-second segments, while the 5-minute EXT intervals are divided into five 1-minute segments. The total number of responses within each segment is calculated and averaged across trials, providing a measure of the development of response within a trial across the experiment.

Mean response rates for the 2-minute fixed-interval component were calculated by dividing each fixed-interval trial into twelve consecutive 10-second segments. The number of responses within each 10-second segment was then summed for each trial. Finally, the mean response rate for each 10-second segment was calculated across trials 2-7 in all components. Similarly, mean response rate for the 5-minute extinction component was calculated by dividing each extinction trial into 5 consecutive 60-second segments. The total number of responses made during each 60-second segment were summed, and then averaged across sessions 32 to 50.

Analyses comparing mean response rates of control and ADHD-model groups showed significant differences in stabilized response behaviour. The ADHD-model group responded approximately twice as much as the control group during both fixed-interval

and extinction components (Figure 3.1), corresponding to hyperactive and inattentive behaviour respectively. Furthermore, the control group spaced their responses with average IRTs > 2-seconds. By contrast, the ADHD-model group was found to produce significantly more responses with short IRTs (< 0.66 seconds), exhibited as rapid responses bursts (Figure 3.2), corresponding to motor impulsivity.

2.1.2 Human experiment: method and results

The multi-FI/EXT human task results outlined in Sagvolden et al. (1998) are of central importance to the hypo-active dopamine theory of ADHD as they extend the animal findings to diagnosed children. While the experimental parameters were kept as close to the animal experiments as possible, there are a few key differences. Most notably, the human task does not investigate stable response behaviour due to the brevity of the experiment; rather, it focuses primarily on behavioural acquisition.

In this experiment, the behaviour of children diagnosed with ADHD was compared to matched controls (see Figure 2.3 for a task diagram). Subjects were 20 boys, eight of which were diagnosed with ADHD and were medication naïve. Each child participated in six consecutive sessions of multi-FI/EXT task disguised as a mechanized game. During the 30-second fixed-interval component, signalled by the game's lights turning on, the first response after a 30-second delay was reinforced with a trinket or coin. Lights on the mechanized game were turned off to signal a 120-second extinction component, during which no reinforcers were delivered. Each session consisted of two components: 1) A 30-second fixed-interval component in which 5 reinforcers were delivered, and 2) a 120-second extinction component.

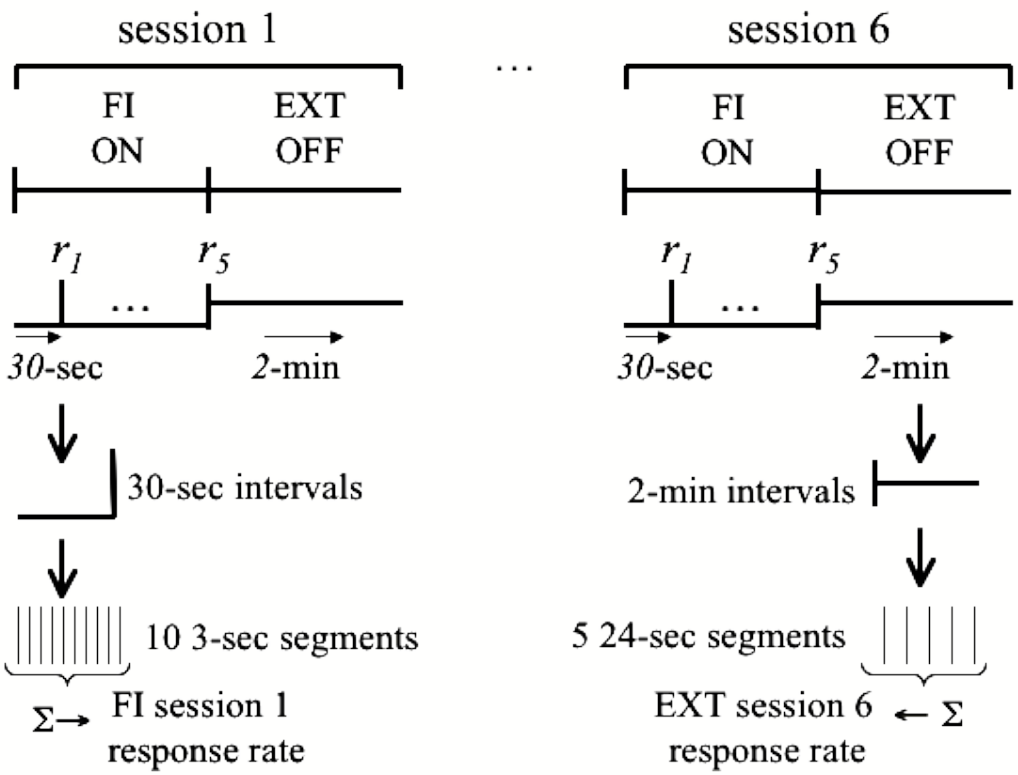

Figure 2.3: *Multi-FI/EXT human task* (Sagvolden et al., 1998). Each of the 6 sessions consist of 2-components: FI→EXT. In each FI component 5 reinforcers are delivered after a fixed delay of 30 seconds. The 30-second FI intervals are divided into ten 3-second segments, while the 2-minute EXT intervals are divided into five 24-second segments. The total number of responses is calculated within each segment and averaged among trials within a component, providing a mean response rate for each session as a function of segment. Segment response rates are summed within each session to provide a total response index for each session.

Analyses reveal results that are qualitatively similar to the animal results outlined above (see (Sagvolden, 2000) and (Sagvolden et al., 2005) for discussions relating animal and human behaviour on a multi-FI/EXT task). When response rates are averaged across all sessions, the ADHD group was observed to respond more frequently during both the fixed-interval and extinction components (Figure 3.9). Furthermore, the control group responded with IRTs > 2-seconds while the ADHD group exhibited response bursts, quantified by IRTs < 0.33. Their results show that by the end of the experiment, the ADHD group exhibits a greater proportion of responses with short IRTs when compared

to the control group (Figure 3.10). Most importantly, the behaviour associated with ADHD was shown to develop as reinforcers were delivered. Control and ADHD group response behaviour was nearly identical at the start of the experiment. However, ADHD group response rate accelerated during both fixed-interval and extinction components across sessions, whereas control group response rate remained constant (Table 3.2 & Figure 3.8).

2.2 Summary

Together, the animal and human results on the multi-FI/EXT task illustrate both acquisition and stabilized behavioural differences between control and ADHD groups. Hyperactivity, motor impulsivity and inattentive behaviour of the ADHD-group were shown to develop as reinforcers were delivered (Sagvolden et al., 1998). Excessive responding during fixed-interval components (i.e. hyperactivity), and extinction components (i.e. attention deficit), along with decreased IRTs (i.e. motor impulsivity) were all shown to remain as stabilized animal ADHD-model behaviour (Sagvolden et al., 1992). These results fall in line with claims that the behaviour associated with ADHD is absent in novel situations (Sleator & Ullmann, 1981) and suggest that behavioural symptoms develop as a result of reinforcers effect on behaviour.

2.3 Model and simulations

The model I propose in this thesis is an application of TDRL to simulations of the multi-FI/EXT tasks outlined previously. I simulate the animal experiment outlined in Sagvolden et al. (1992) to explore stable response behaviour, and the human experiment discussed in Sagvolden et al. (1998) to explore acquisition. In the following section I provide an outline of my model and task simulations.

2.3.1 Task Simulation

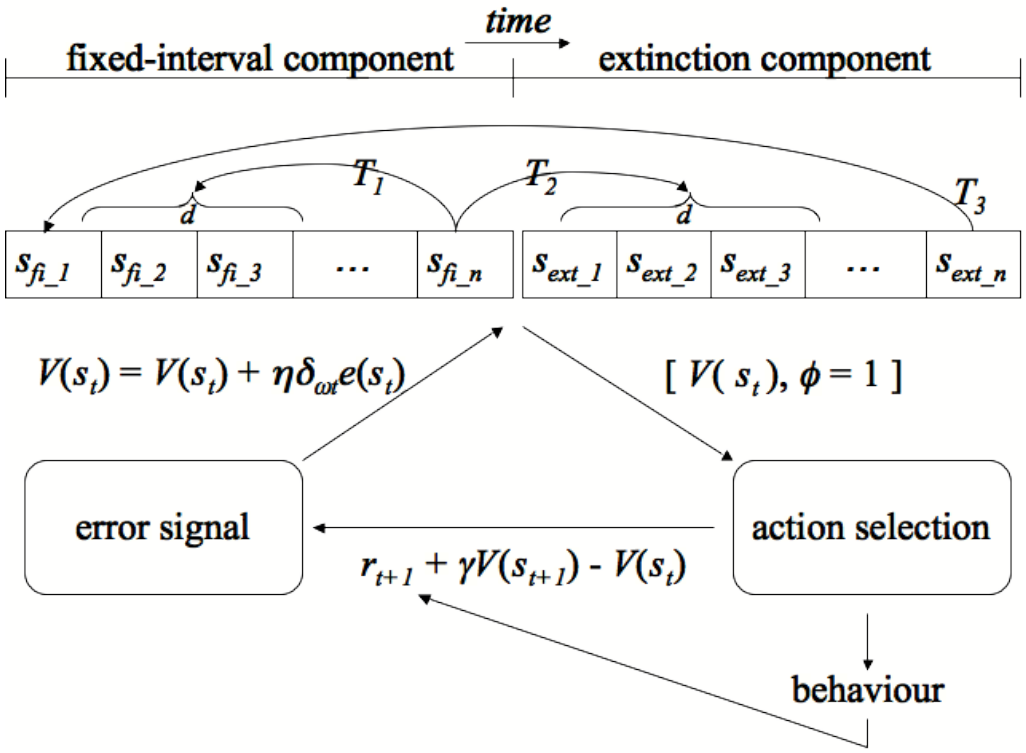

Figure 2.4: *Model architecture and dynamics*. *State-space*: Each state, s_t, represents an equal proportion of elapsed time until the end of a component trial. Hence, state $s_{fi\_k} = s_{fi\_l} = s_{ext\_m} = 1/n$ percent of trial duration. The state values, $V(s_t)$, represent the expected reward at that particular moment in time. *State-transitions*: Transitions represent the passing of time, from one moment in a trial to the next. Hence, transitions are largely independent of the agent's behaviour (e.g. $s_{fi\_1} \rightarrow s_{fi\_2} \rightarrow \ldots$). During a fixed-interval trial, the first response after the required delay is defined as a terminal state with $r_T = 1$, forcing the agent along transition T_1 to begin a new fixed-interval trial starting at state s_d. When the last terminal state of a fixed-interval component is encountered, the agent follows transition T_2 into the extinction state space starting at state s_d. Until the simulation is complete, transition T_3 leads the agent out of the last extinction state into s_{fi_1} (because there is no reward reception delay following an extinction trial) to begin a new fixed-interval trial. However, if the experiment is complete the last extinction state is defined as a terminal state with $r_T = 0$. *Action selection*: At each state a "softmax" function, defined by equation(2.1), is used to probabilistically select between $V(s_t)$ and a constant response threshold, $\phi = 1$. If $V(s_t)$ is selected, then the agent makes a "response", otherwise no response is made. Each response incurs a small penalty $r_c = -0.05$. *Error signal and value update*: After each action selection, an error signal, δ_t, is calculated according to equation

(1.2). A small amount of noise is added to the error signal, which is then scaled according to equation (2.2). Finally, the value function for $V(s_t)$ is updated according to equation (1.5), using $\delta_{\omega t}$ as the error signal.

Figure 2.4 illustrates the structure and learning dynamics of the multi-FI/EXT task simulation. A multi-FI/EXT task require subjects to discriminate between times when a response is beneficial and those that are not. The environment state-space defines what a TDRL model is capable of learning; hence, I define a state-space in which each state represents a fixed period of time. Given this state-space structure, state values represent the value of responding at any given moment during the experiment. As the model interacts with the environment it can learn higher state values for times when responses are likely to be profitable, and lower values when they are not.

The model traverses a finite set of states, one subset representing a fixed-interval component trial, the other an extinction component trial. Gallistel & Gibbon (2000) have demonstrated that a timing model featuring timescale invariance can account for a wide range of conditioning phenomena, including behaviour on a fixed-interval of reward delivery. Hence, my simulations use a state-space with timescale invariance in which each state represents an equal proportion of time until the end of the trial. As such, an equal number of states represent the time-course of both fixed-interval and extinction trials.

There is only one possible transition into each state, simulating the passage of time, independent of the model's actions. This implies that my simulation satisfies the requirements of a finite Markov decision process, an important assumption of any TDRL algorithm (Sutton & Barto, 1998). A state signal has the Markov property if the environment's response at time $t+1$ depends only on the state, s_t, and action, a_t, at time t. An environment is an MDP if and only if:

$$Pr\{ s_{t+1} = s', r_{t+1} = r \mid s_t, a_t, r_t, s_{t-1}, a_{t-1}, r_{t-1}, \ldots, r_1, s_0, a_0 \}$$
$$=$$
$$Pr\{ s_{t+1} = s', r_{t+1} = r \mid s_t, a_t \}$$

for all s', r, s_t and a_t. Hence, only the current state and action need be known in order to predict the next state and corresponding reward; any state history larger than this is redundant. My task simulation meets this criterion since there is only a single transition into each state, and actions (i.e. responses) are always met with the same reward for a given state.

I simulate the time required to receive a reinforcer and return to the task at hand by selecting a positive delay time, d, pulled from a normal distribution, $N(10, 40)$, at the start of each learning trial. This distribution defines the set of possible starting states for the simulation. Starting in state s_d, the model decides whether it should respond or not, then transitions into state s_{d+1}. While traversing the fixed-interval state space, the first decision to respond after the required fixed-interval delivers a reinforcer, $r = 1$, and terminates the learning trial. The model then starts the next learning trial by selecting a new starting state, s_d', from the appropriate subset of fixed-interval or extinction states. Once a fixed-interval component is complete, the model starts the next learning trial in the extinction state space. After traversing the extinction state space, the final extinction state transitions into either the first fixed-interval state, simulating a new fixed-interval component, or if the simulated session is complete, into a terminal state with $r = 0$.

The simulation structures are as close to the original experiments as possible. I simulate the animal experiment described in Sagvolden et al. (1992) with 58 sessions, each session consisting of four parts: 1) A fixed-interval component in which a maximum of seven reinforcers are delivered, 2) an extinction component with no reinforcer delivery, 3) a second fixed-interval component with the same properties as the first, and 4) a final extinction component to complete the session. The state space is made up of 240 states, 120 for fixed-interval and 120 for extinction trials. The model's mean response rate for the fixed-interval component includes data from trials 2-7 of all 58 sessions, while the extinction component mean response rate includes all trials from sessions 32-50 as described in Sagvolden et al. (1992). Since the state-space has no notion of absolute time, only temporal proportion, I define IRTs as either short or long. Short IRTs are defined as

consecutive non-terminal states, $s_t \rightarrow s_{t+1}$, in which a response was made at both s_t and s_{t+1}. All states in which a response was made that do not meet this criterion are considered long IRTs.

The human experiment described in Sagvolden et al. (1998) is simulated as 6 sessions, each consisting of 2 parts: 1) a fixed-interval component in which five reinforcers are delivered, 2) an extinction component where no reinforcers are delivered. The state-space consists of 480 states in total, 240 representing a fixed-interval trial and 240 representing an extinction trial. The model's mean response rate for both fixed-interval and extinction components include all trials from all sessions. IRTs are defined in the same way as in the animal task simulation.

2.3.2 Model

My proposed model operates on the environment previously outlined, attempting to learn the optimal response behaviour. After each transition, $s_t \rightarrow s_{t+1}$, a reward prediction error, δ_t, is calculated as defined by equation (1.2). A small amount of noise, ε_δ, sampled from a normal distribution *N(0, 0.1)*, is added to δ_t to account for the imprecision of a biological system. State values for all previously encountered states are then updated according to equation (1.5).

Action selection was determined according to a 'Softmax' probability function (Egelman, Person, & Montague, 1998; Sutton & Barto, 1998). At each state, the model makes a binary decision whether or not to make a single response as defined by:

$$\mathrm{P}(\text{ response } k \text{ selected }) = \frac{e^{\varphi_t^k / \tau}}{\sum_{k=1}^{2} e^{\varphi_t^k / \tau}} \qquad (2.1)$$

where φ^k_t corresponds to the value associated with responding at time t, $V(s_t)$, or a constant response threshold, $\phi = 1$. τ is a 'temperature' parameter that controls the degree of exploration. To discourage the model from adopting a strategy of simply responding at each state, each response incurs a small cost, $r = r_c = -0.05$, which is factored into the prediction error calculation for that state.

The procedure for investigating the animal and human simulations was identical, though target optimization data was different. For each simulated experiment the proposed model was run 30 times, simulating data from 30 subjects. I began by searching for set a of η, γ, λ, and τ parameter values that minimized the discrepancy between the model's behaviour and that of the control group's empirical data. The resulting optimal parameter set defines the "control" model. This "control" model provides the base from which my investigation into the impact of prediction error size is rooted. As was noted earlier, the animal experiment outlined in Sagvolden et al. (1992) was primarily concerned with stabilized response behaviour. Hence, I used the mean response rate across sessions from fixed-interval and extinction components, as the optimization target for the animal simulation. The human experiment outlined in Sagvolden et al. (1998) focused on behavioural acquisition. Therefore, I used the total number of responses within each session's mean response rate as the optimization target for the human simulation optimization. Total session response rates were calculated by summing the segmented response rates for each session's mean response rate, providing an index of response rate across sessions for each component.

Values for η, γ, λ, and τ parameters were determined separately for animal and human simulations. I used a constrained gradient descent algorithm to find an optimal set of parameters that minimize the discrepancy, measured as the sum of squared error, between target empirical data and the model's response behaviour. Possible values for η, γ, λ, and τ were all bound between 0 and 1, while all other parameter values were locked. I seeded the search algorithm with a random set of starting parameter values from which it began to search for an optimal fit. To help avoid the possibility of finding a sub-optimal local

minimum solution, this process was repeated 100 times for each task simulation using a random set of starting parameter values each time.

Once the control model had been defined, I investigated the implications of hyper- and hypo-active midbrain dopamine systems. This was done by locking the "control" model's parameters and introducing prediction error scaling parameters, $\omega_{\delta+}$ *and* $\omega_{\delta-}$. After the noise term, ε_{δ}, has been added to δ_t, the prediction error is scaled by:

$$\delta_{wt} = \begin{cases} \omega_{\delta+}(\delta_t) & \text{if } \delta_t > 0 \\ \omega_{\delta-}(\delta_t) & \text{if } \delta_t < 0 \end{cases} \tag{2.2}$$

where $\omega_{\delta+} = \omega_{\delta-} = 1$ is considered a "normal" scaling factor ($0 \leq (\omega_{\delta+}, \omega_{\delta-}) \leq 2$). State values were then updated according to equation (1.5), using $\delta_{\omega t}$ as the reward prediction error value. The hyper-active dopamine theory corresponds to scaling factors of $\omega_{\delta+} = \omega_{\delta-} > 1$, while the hypo-active dopamine theory maps onto scaling factors of $\omega_{\delta+} = \omega_{\delta-} < 1$. I quantify the impact of scaling the prediction error along a continuum of values by the discrepancy, measured as the sum of squared error, between the ADHD group's empirical target data and model's corresponding behaviour.

Finally, I conducted a second parameter search aimed at finding a set of parameter values that minimize the discrepancy between the model's behaviour and that of the hyperactive group. This "ADHD fit" model facilitated an investigation of possible roles played by other parameters in ADHD.

3 Results

3.1 Animal simulation results

Model	***SSE***	η	γ	λ	τ	$\omega_{\delta+}$	$\omega_{\delta-}$
Control	0.64	0.50	0.99	0.50	0.50	1.00	1.00
Hyper-/Hypo-	66.64	0.50	0.99	0.50	0.50	1.70	1.70
Asymmetrical	10.00	0.50	0.99	0.50	0.50	1.78	1.32
ADHD Fit	19.66	0.32	0.99	0.99	1.00	1.00	1.00

Table 3.1: *Measure of fit and parameter values for animal models*. Sum of squares error, (SSE), quantifies the discrepancy between mean response rates of the model and empirical data: "Control" SSE = (model – control group)^2, "Hyper/Hypo" = (model – ADHD group) ^{2}, "Asymmetrical" SSE = (model – ADHD group) ^{2}, "ADHD Fit" = (model – ADHD group) ^{2}. Light grey cells indicate parameters that were allowed to vary while exploring a given model, whereas dark grey cell values are locked.

Parameter values for the different models are outlined in Table 3.1, while Figure 3.1 illustrates mean response rates for those models along side target empirical data from animal groups as reported in Sagvolden et al. (1992). As can be seen by inspection, the "control" model's mean response rate closely matches that of the animal control group for both fixed-interval and extinction components (SSE is included in Table 3.1). This is unsurprising given that the model's parameters were optimized to fit the animal control group mean response rates, but it demonstrates that the model can replicate "normal" response rate behaviour. Figure 3.2 illustrates the mean proportion of impulsive responses during the fixed-interval component for both animals and models. Note that empirical IRT data was not included in the optimization search. Hence, the model's IRT behaviour is not due to finding a set of parameters that match the empirical control group's IRT distribution; rather, it is an emergent property of the model itself. The "control" model's proportion of short IRTs is comparable to that of the animal control group. Thus, I conclude that the "control" model exhibits "normal" levels of hyperactive, impulsive and inattentive behaviour as quantified by the animal control group in Sagvolden et al. (1992). Furthermore, an exploration of the parameter values used to

define the control model shows that its behaviour is robust under a range of parameter values. Figure 3.3 shows that learning rate, discount factor and decay rate can hold a wide range of values without increasing model error, while response threshold appears to have an optimal basin.

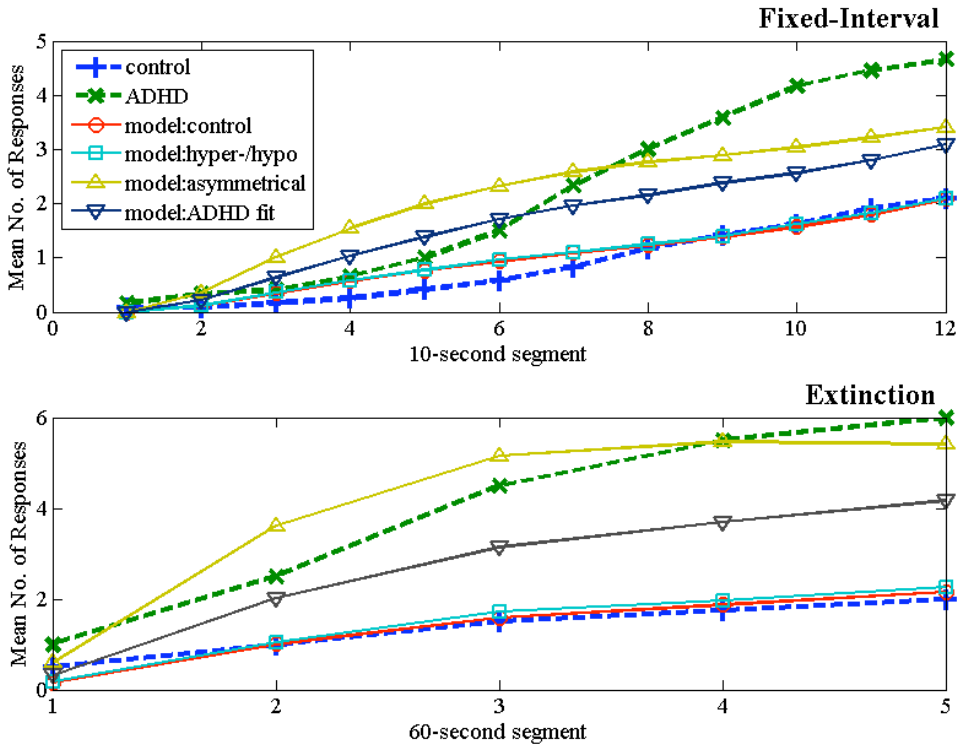

Figure 3.1: *Stabilized animal and model response development within multi-FI/EXT task trials*. Empirical and simulated mean response rates for fixed-interval (above) and extinction (below) components. Mean response rates of control and animal ADHD-model groups as reported in Sagvolden et al. (1992), and mean response rates of the model with parameter values as outlined in Table 3.1. *x-axis*: Each trial is divided into twelve 10-second segments (fixed-interval) or five 60-second segments (extinction) for which the number of responses is summed. *y-axis*: Mean number of responses for each segment across trials 2-7 of each fixed-interval component, and all extinction component trials from sessions 32-50.

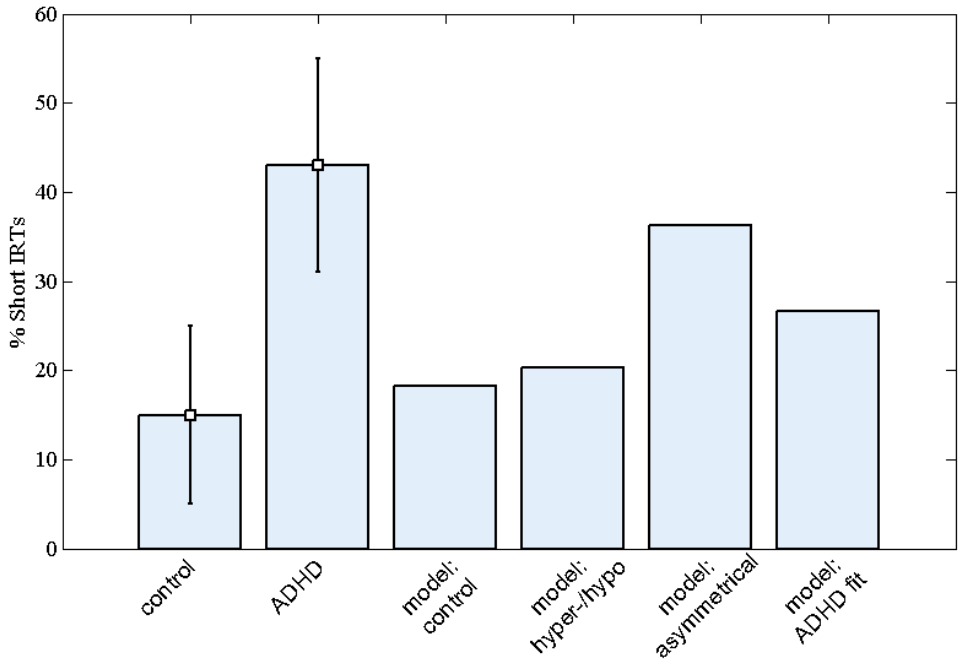

Figure 3.2: *Stabilized animal and model inter-response time during the multi-FI/EXT task*. Empirical and simulated proportion of responses with short IRTs. Impulsive IRTs after behavioural stabilization for animal control and ADHD-model groups as reported by Sagvolden (2000) (modified after Sagvolden et al. (1992)), and of the model with parameter values as outlined in Table 3.1. *y-axis*: Short IRTs in the model are defined as consecutive states in which a response was made. Short IRTs for animals are defined as responses with IRTs < 0.67-seconds.

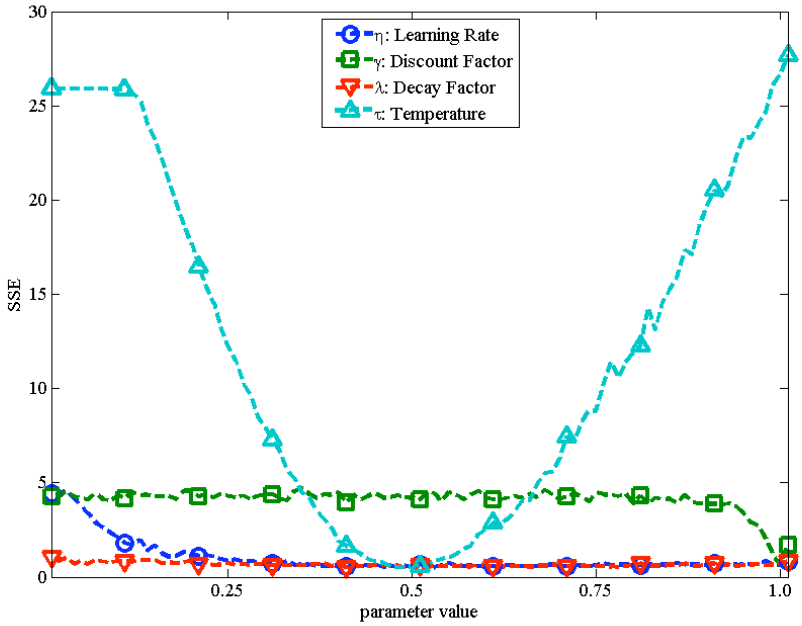

Figure 3.3: *Animal control model parameter robustness*. To provide a measure of model robustness with respect to optimal parameter values, each parameter was varied through its range of possible values in isolation. All other parameters are locked at "control" model values while a single parameter value is varied. *x-axis*: parameter value. *y-axis*: Discrepancy between the model's response rate and that of animal controls.

As discussed previously, I investigated the hyper- and hypo-active dopamine theories of ADHD by scaling the prediction error size through a continuum of values above and below a "normal" scaling factors of $\omega_{\delta+} = \omega_{\delta-} = 1$. The hyper-active dopamine theory of ADHD proposed by Grace (2001) predicts improved agreement between model and animal ADHD-model group response rates for scaling factors above 1, while the hypo-active theory proposed in Sagvolden et al. (2005) predicts increased agreement for scaling factors below 1. Figure 3.4 depicts the discrepancy between ADHD-model group and the "control" model's mean response rates while varying prediction error scaling factors across a range of values ($0 \leq \omega_{\delta+} = \omega_{\delta-} \leq 2$). There is no notable improvement in the disparity between simulated and empirical behaviour for either hyper-active ($\omega_{\delta+} = \omega_{\delta-} > 1$) or hypo-active ($\omega_{\delta+} = \omega_{\delta-} < 1$) scaling factors. Furthermore, an analysis of the "hyper-/hypo" model behaviour exhibited by the optimal fit (denoted by the red circle in Figure 3.4) reveals behaviour that resembles that of the control group more so than that of

the SHR group (Figure 3.1 & Figure 3.2). Hence, neither abnormally large nor abnormally small prediction errors appear to be capable of producing ADHD-like behaviour as suggested by the hyper- and hypo-active dopamine theories of ADHD. A more detailed analysis of each parameters ability to induce ADHD behaviour reveals that the only parameter capable of significantly improving the model's fit to ADHD behaviour is response temperature (see Figure 3.5). Temperature's role in ADHD behaviour will be discussed shortly.

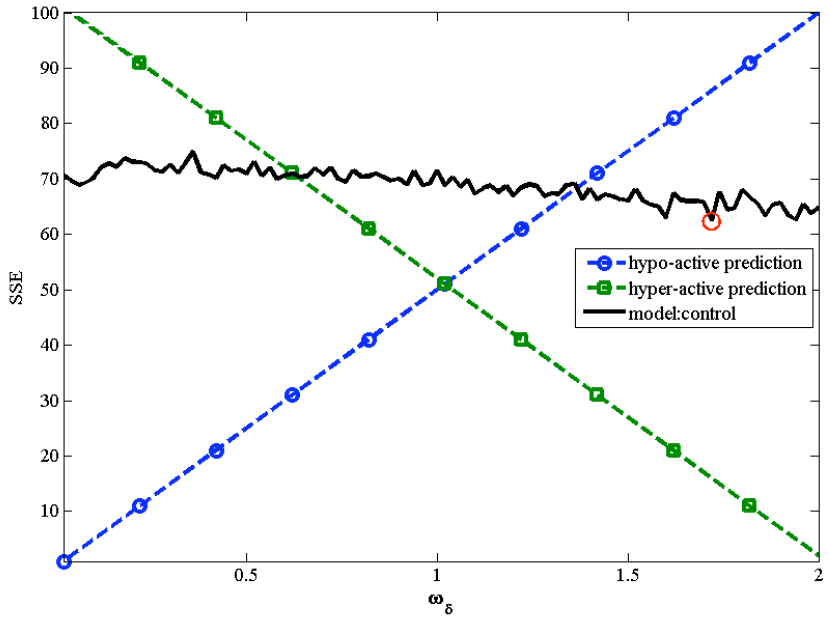

Figure 3.4: *Theoretically predicted and simulated animal response behaviour as a function of prediction error size*. The discrepancy between animal ADHD-model group and simulated mean response rates as a function of prediction error size (solid line). The hyper-active dopamine theory predicts ADHD-like behaviour to emerge with abnormally large prediction errors (green dashed line), while the hypo-active theory predicts such behaviour will depend on abnormally small prediction errors (blue dashed line). Note that hyper-/hypo-active predictions diagrammed here are a simplification. Neither theory predicts a linear relationship between DA signals and ADHD behaviour. The red circle indicates the scaling factor that provides the best fit to animal ADHD-model group behaviour, defining the "hyper-/hypo" model. *x-axis*: Scaling factors of equal size for positive and negative prediction errors, $\omega_\delta = (\omega_{\delta+} = \omega_{\delta-})$. *y-axis*: Sum of squares difference relative to animal ADHD-model group mean response rate.

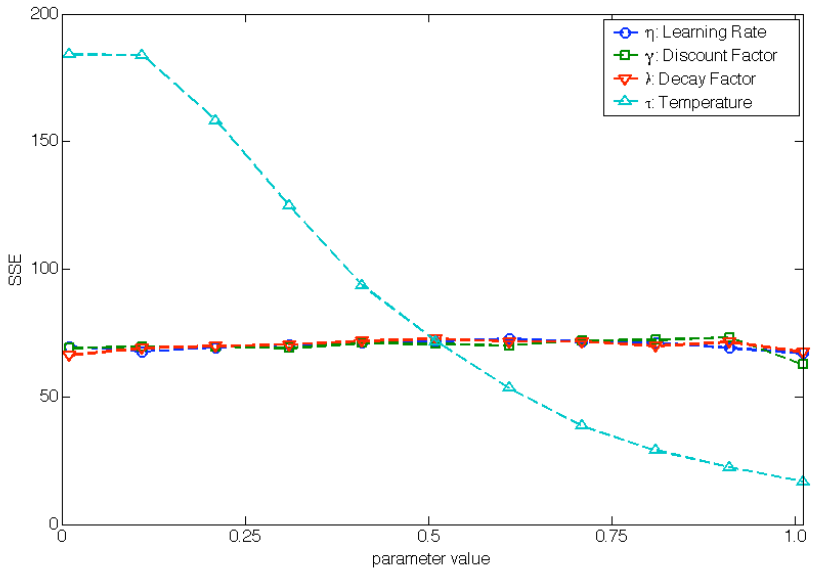

Figure 3.5: *Animal ADHD model isolated parameter manipulation:* To investigate the how each parameter may contribute to ADHD-like behaviour, each parameter was varied through its range of possible values and compared to behaviour of the animal model of ADHD. All other parameters are locked at "control" model values while a single parameter value is varied. *x-axis*: parameter value. *y-axis*: Discrepancy between the model's response rate and that of ADHD animals.

While a formal investigation of the implications of abnormally large and small prediction errors revealed that neither provides a satisfactory mechanism for producing ADHD-like behaviours, a model-based examination allows me to explore a wider range of dopaminergic dysfunction. To this effect, I varied the prediction error scaling factors, $\omega_{\delta+}$ and $\omega_{\delta-}$, independently while all other "control" model parameter values were locked. Hence, I was able to explore the implication of large/small positive prediction errors independently of large/small negative prediction errors. Simulations were run comparing model response rate to both animal control and ADHD-model groups (Figure 3.6, left panel). Focusing on model behaviour relative to control group behaviour note that: 1) the optimal fits are found along the diagonal, where $\omega_{\delta+} \approx \omega_{\delta-}$, and 2) the optimal fits extend above and below "normal" scaling factors of $\omega_{\delta+} = \omega_{\delta-} = 1$. A different pattern of results emerges when comparing model and animal ADHD-model group behaviour. Figure 3.6

(right panel) shows that optimal fits to animal ADHD-model behaviour are found above the diagonal, where $\omega_{\delta+} > \omega_{\delta-}$; but again, optimal fits extend across a range of scaling factors both above and below "normal" scaling factors of $\omega_{\delta+} = \omega_{\delta-} = 1$.

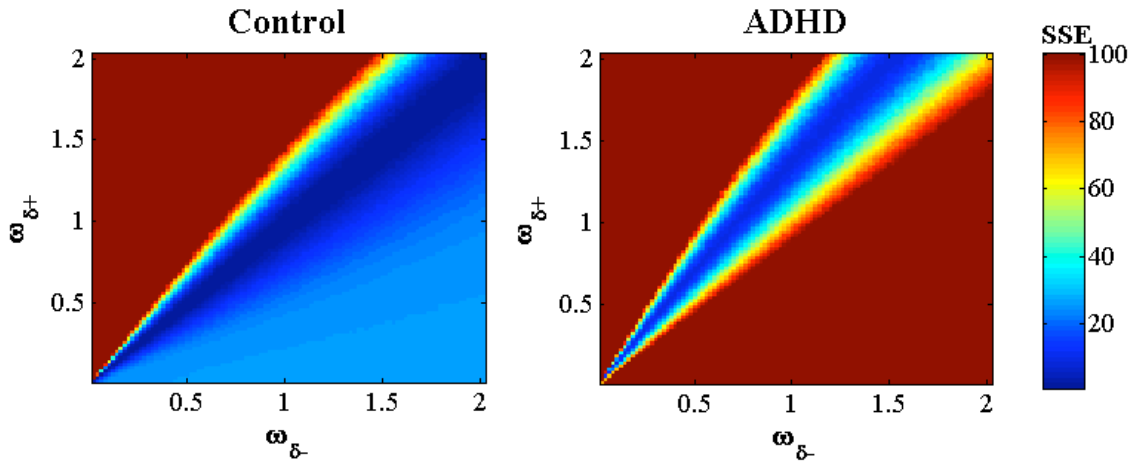

Figure 3.6: *Impact of positive and negative prediction error size.* Simulated behaviour compared to animal control (left) and ADHD-model (right) group mean response rates while varying the size of positive and negative prediction errors independently. *x-axis*: size of negative prediction error ($\omega_{\delta-}$). *y-axis*: size of positive prediction error ($\omega_{\delta+}$). *Colour*: Sum of squared error between simulated and empirical mean response rates. Blue indicates lower error while red indicates higher error.

A common factor observed among optimal fits for both animal group comparisons is the ratio $R_\omega = \omega_{\delta+}/\omega_{\delta-}$. The relationship between the prediction error ratio R_ω, and its impact on behaviour is illustrated in Figure 3.7. Animal control group behaviour is best fit by models with a prediction error ratio of $R_\omega \approx 1$. Interestingly, animal ADHD-model group behaviour is best fit by models with $R_\omega \approx 1.2$, which I refer to as "asymmetrical" models (see Table 3.1 for optimal parameter values). For this model, the mean response rates of both fixed-interval and extinction components (Figure 3.1), as well as the IRT distribution (Figure 3.4) of the "asymmetrical" model better resembles the animal ADHD-model group. Hence, it appears that the prediction error ratio, R_ω, can be varied such that the model's response behaviour matches that of the SHR group as reported in Sagvolden et al. (1992). Specifically, an asymmetrical prediction error, where positive

error signals are larger than negative signals, produced the hyperactive, impulsive and inattentive behaviours associated with ADHD. It is this asymmetrical ratio between positive and negative error signals, not the absolute signal size that appears crucial for ADHD behaviour to emerge.

A final exploration was conducted by unlocking the η, γ, λ, and τ parameters and running the optimization algorithm against the animal ADHD-model mean response rate. The search was able to find a set of parameter values, independent of prediction error size, such that the model's mean response rate matched that of the animal ADHD-model group reasonably well (Table 3.1). I refer to this as the "ADHD fit" model. The "ADHD fit" model's mean response rates for both components (Figure 3.1) and the IRT distribution (Figure 3.2) shows that it does indeed reproduce the hyperactive, impulsive, and inattentive behaviour of the animal ADHD-model group. An investigation into the "ADHD fit" parameter values shows that this behaviour is largely the result of emphasizing response exploration via a high temperature value, τ (see Figure 3.3). This model is more likely to choose actions independent of learned values since the high response selection temperature will tend to reduce the discrepancy between response choice values.

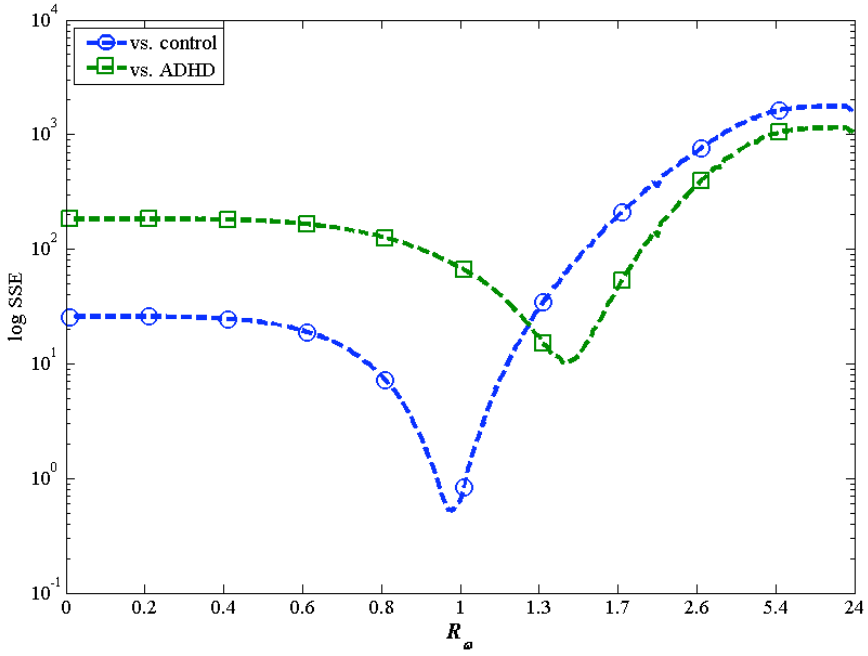

Figure 3.7: *Impact of prediction error ratio on animal model behaviour*. Simulated behaviour compared to animal control (blue dashed line) and ADHD-model (green dashed line) group mean response rates relative to the ratio of positive and negative prediction error size, $R_\omega = \omega_{\delta+} / \omega_{\delta-}$. Ratios were calculated for all values illustrated in Figure 3.6 ($0 \leq \omega_{\delta+}$, $\omega_{\delta-} \leq 2$), sorted, and pooled into 200 bins. *x-axis*: mean ratio of each bin. *y-axis*: Log sum of squared error between simulated mean response rate averaged within each ratio bin, and animal ADHD-model group mean response rate.

My simulation of the animal experiment outlined in Sagvolden et al. (1992) reveals two possible mechanisms driving the hyperactive, impulsive and inattentive behaviour observed in the animal ADHD group. One possible mechanism responsible for these behaviours is characterized by asymmetrical positive and negative prediction errors such that the prediction error ratio, $R_\omega = \omega_{\delta+}/\omega_{\delta-} > 1$. A second possible mechanism giving rise to ADHD-like behaviour is an abnormally high rate of response exploration. Here, learned values are nearly inconsequential in determining response behaviour. Hence, decisions are essentially random, which translates into an increase in response rate. In order to dissociate these two potential mechanisms underlying ADHD I proceed to my human experiment simulation in order to investigate behavioural acquisition.

3.2 Human simulation results

Subject	***SSE***	***FI slope***	***EXT slope***	η	γ	λ	τ	$\omega_{\delta+}$	$\omega_{\delta-}$
Children									
Control	-	-0.09	-3.1	-	-	-	-	-	-
ADHD	-	7.5	3.5	-	-	-	-	-	-
Model									
Control	1412	0.3	-1.65	0.50	0.99	0.95	0.63	1.00	1.00
Hyper-/Hypo-	5000	0.3	0.04	0.50	0.99	0.95	0.63	0.00	0.00
Asymmetrical	1832	5.1	1.6	0.50	0.99	0.95	0.63	0.68	0.42
ADHD Fit	2191	-0.3	-0.82	0.4	0.94	0.90	0.94	1.00	1.00

Table 3.2: *Measure of fit and parameter values for human models*. Sum of squares error, (SSE), quantifies the discrepancy between session response rates of the model and empirical data: "Control" SSE = (model – control group)^2, "Hyper/Hypo" = (model – ADHD group) ^{2}, "Asymmetrical" SSE = (model – ADHD group) ^{2}, "ADHD Fit" = (model – ADHD group) ^{2}. Linear regression slope coefficients index the response rate change across sessions for both fixed-interval and extinction components. Light grey cells indicate parameters that were allowed to vary while exploring a given model, whereas dark grey cell values are locked.

The animal model simulation was explored in terms of stabilized mean response rate averaged across sessions as reported in Sagvolden et al. (1992). Since these data do not include response rates as a function of session, it is necessary to include a simulation of the human experiments outlined in Sagvolden et al. (1998) to explore behavioural acquisition as a function of reinforcement. Hence, the response rate across sessions was optimized to define a human "control" model in order to emphasize behavioural changes rather than stabilized response rates. Parameter values for the different models are outlined in Table 3.2. Figure 3.8 depicts the session response rate of the "control" model defined by the optimal set of parameter values arrived at by the optimization algorithm. The regression slope coefficients for the "control" model and empirical control group exhibit the same general trend (slope Table 3.2). Fixed-interval response rate remains consistent across sessions, whereas extinction response rate decreases across sessions, showing that the search algorithm was indeed able to find a reasonable optimization solution. Despite using only session response rate for optimization, the "control" model

exhibits mean response (Figure 3.9) and IRT (Figure 3.10) behaviour matching that of the control group. Furthermore, an investigation of the control model's parameter values show a similar robustness to what was demonstrated for the animal model simulations (see Figure 3.11).

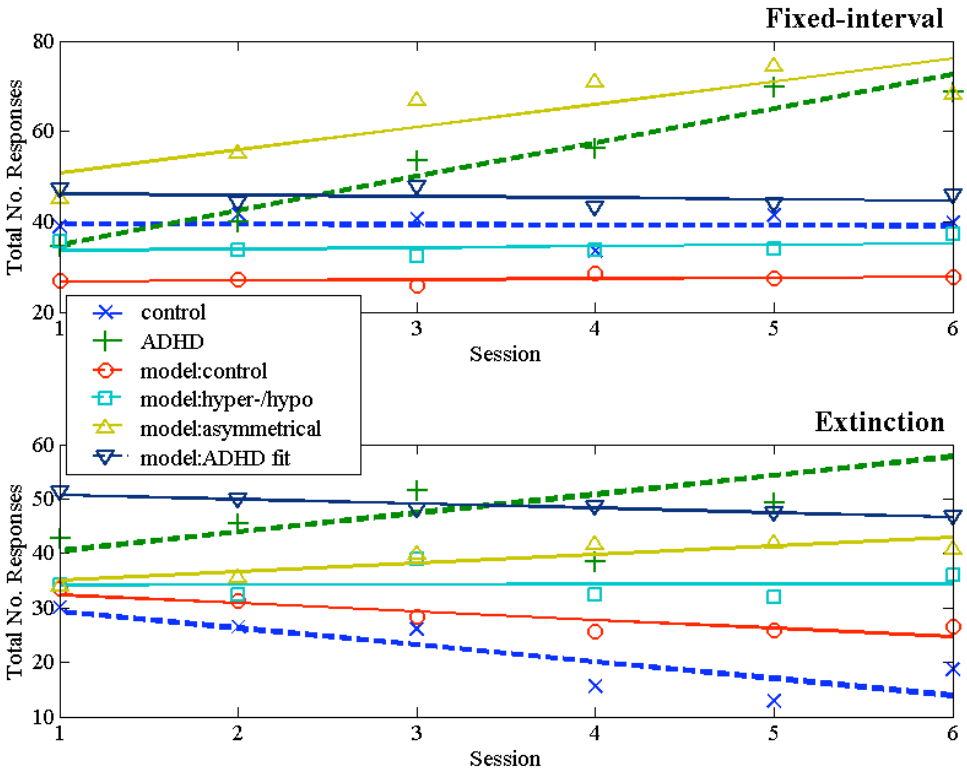

Figure 3.8: *Human response development across multi-FI/EXT task sessions*. Session response rates (markers), and linear regressions (lines) for fixed-interval (above) and extinction (below) components calculated from Sagvolden et al. (1998). *x-axis*: experiment session. *y-axis*: Session response rate, calculated by summing all segment response rates for each session's mean response rate.

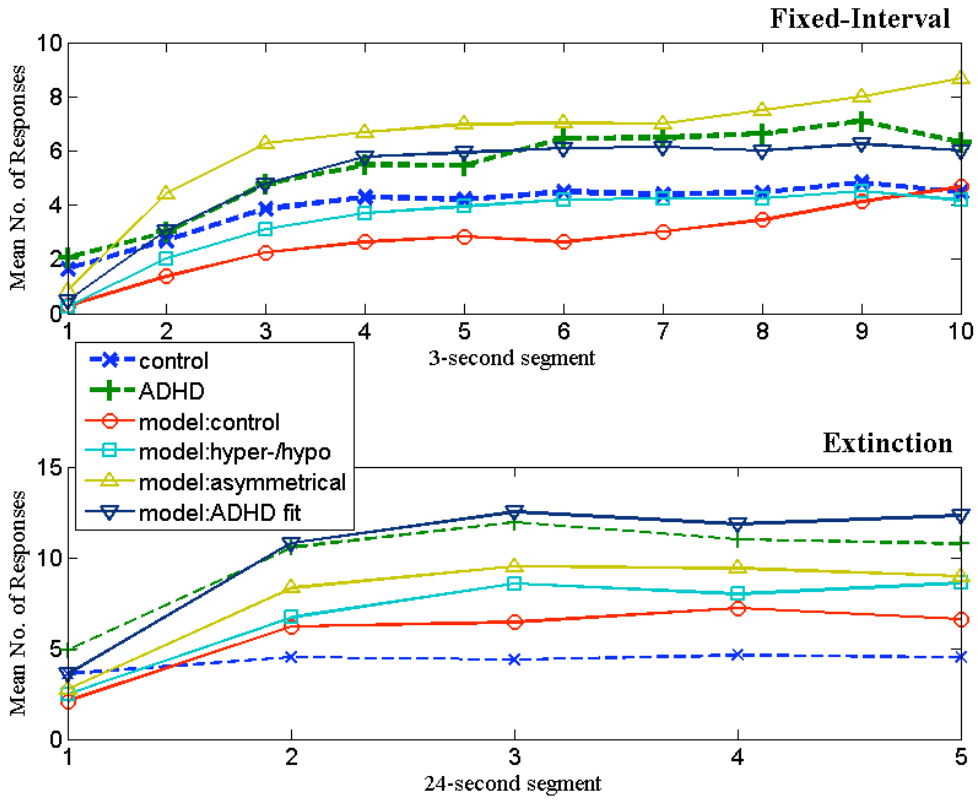

Figure 3.9: *Human response development within multi-FI/EXT task trials*. Empirical and simulated mean response rates for fixed-interval (above) and extinction (below) components. Mean response rates of control and ADHD groups calculated from Sagvolden et al. (1998), and mean response rates of the model with parameter values as outlined in Table 3.2. *x-axis*: Each trial is divided into ten 3-second segments (fixed-interval) or five 24-second segments (extinction) for which the number of responses is summed. *y-axis*: Mean number of responses for each segment across all trials.

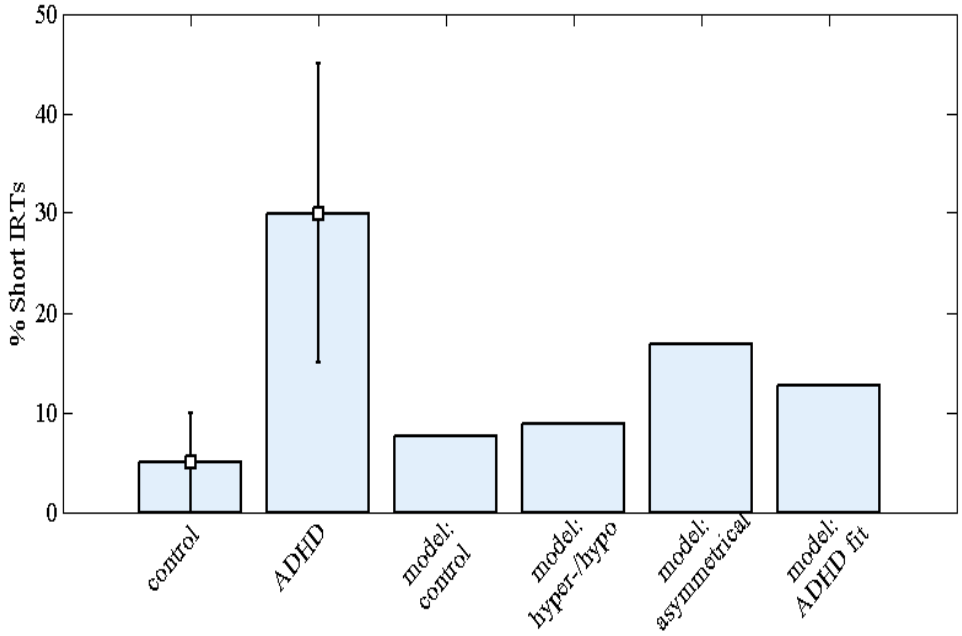

Figure 3.10: *Human inter-response time during the multi-FI/EXT task*. Empirical and simulated proportion of responses with short IRTs. Impulsive IRTs in the last fixed-interval component for control and ADHD groups as reported in Sagvolden et al. (1998), and of the model with parameter values as outlined in Table 3.2. *y-axis*: Short IRTs in the model are defined as consecutive states in which a response was made. Short IRTs for children are defined as responses with IRTs < 0.33-seconds.

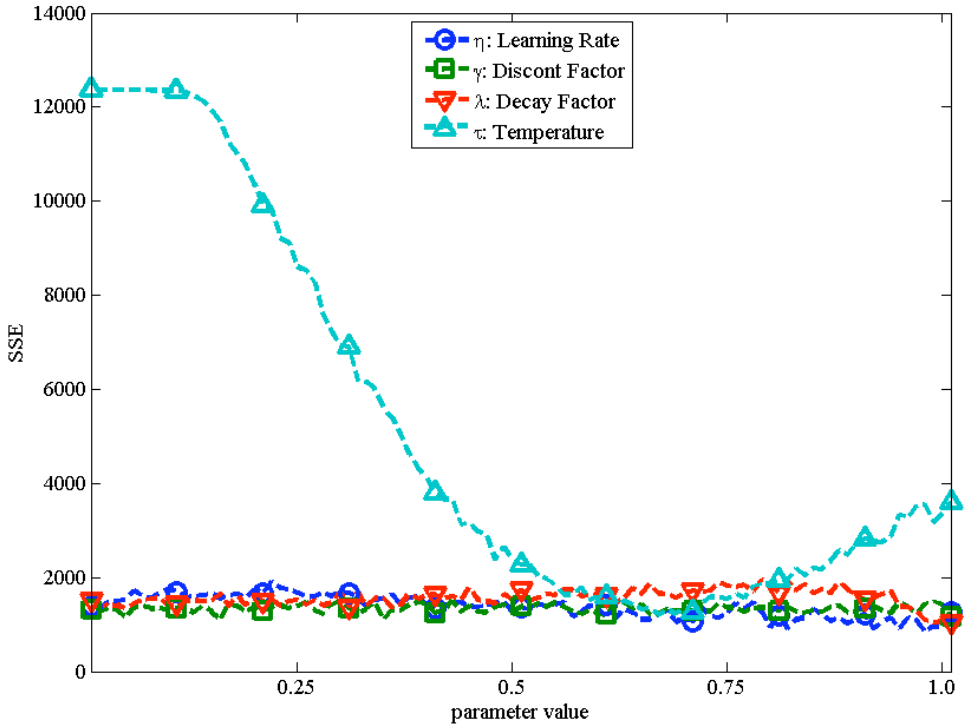

Figure 3.11: *Human control model parameter robustness*. To provide a measure of model robustness with respect to optimal parameter values, each parameter is varied through its range of possible values in isolation. All other parameters are locked at "control" model values while a single parameter value is varied. *x-axis*: parameter value. *y-axis*: Discrepancy between the model's session rate and that of human controls.

Following the same line of investigation as in the animal simulations, I arrive at qualitatively similar "hyper-/hypo", "asymmetrical", and "ADHD fit" models for the human task simulation. Scaling the size of the prediction error to be abnormally large or small did not produce a reasonable fit to the empirical ADHD session response rates (Figure 3.12). The "hyper-/hypo" model defined by the prediction error scaling factors providing the best fit to ADHD session response rate (denoted by the red circle in Figure 3.12) exhibits neither mean response rates (Figure 3.9), nor the IRTs (Figure 3.10) matching ADHD group behaviour. While ADHD group shows a sharp increase in response rate across sessions (Figure 3.8, slope Table 3.2), the "hyper-/hypo" model exhibits almost no change in either fixed-interval or extinction response rates (Figure 3.8,

slope Table 3.2). Again, as was observed with the animal simulations, investigating the role of each parameter in isolation reveals that only the temperature parameter is capable of inducing behaviour more like that of children with ADHD (see Figure 3.13). In particular, I note that decreasing the discount factor, learning rate or the decay factor, as suggested by the hypo-active theory of ADHD, do not produce a significant improvement in the model's fit to ADHD response behaviour.

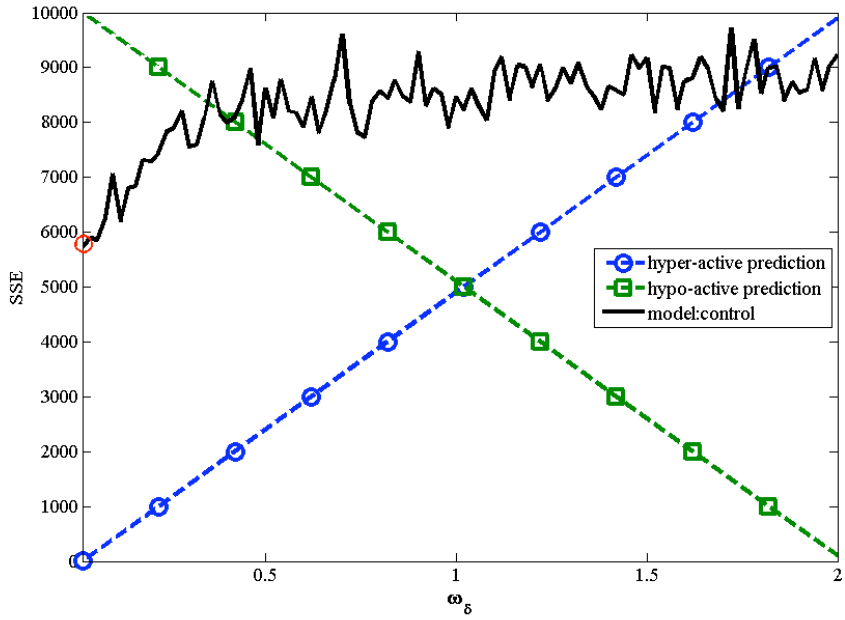

Figure 3.12: *Theoretically predicted and simulated human response behaviour*. The discrepancy between ADHD group and simulated session response rates as a function of prediction error size (solid line). The hyper-active dopamine theory predicts ADHD-like behaviour to emerge with abnormally large prediction errors (green dashed line), while the hypo-active theory predicts such behaviour will depend on abnormally small prediction errors (blue dashed line). The red circle indicates the scaling factor that provides the best fit to ADHD group behaviour, defining the "hyper-/hypo" model. *x-axis*: Scaling factors of equal size for positive and negative prediction errors, $\omega_\delta = (\omega_{\delta+} = \omega_{\delta-})$. *y-axis*: Sum of squares error relative to ADHD group session response rate.

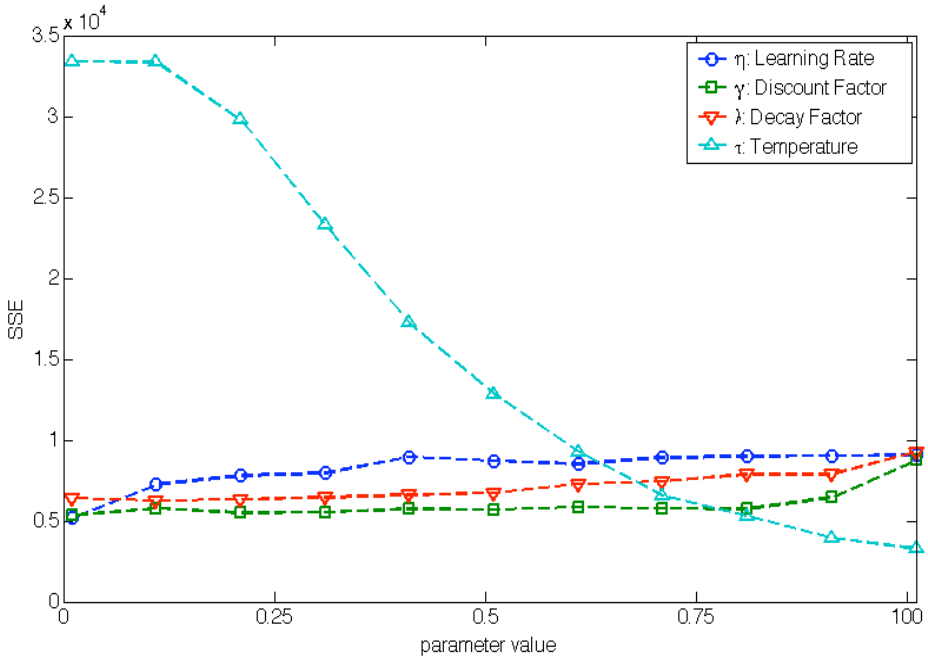

Figure 3.13: *Human ADHD model isolated parameter manipulation:* To investigate the how each parameter may contribute to ADHD-like behaviour, each parameter was varied through its range of possible values and compared to behaviour of the children with ADHD. All other parameters are locked at "control" model values while a single parameter value is varied. *x-axis*: parameter value. *y-axis*: Discrepancy between the model's response rate and that of ADHD children.

However, as was observed for the animal simulations, scaling positive and negative prediction errors independently to produce an "asymmetrical" model for the human simulation shows a drastic improvement in fit (SSE Table 3.2). Figure 3.14 depicts the discrepancy between model and empirical session response rate as a function of positive and negative prediction error size. The empirical session response rates of the control group is best fit along the diagonal where positive and negative scaling factors are balanced, extending above and below the "normal" scaling factor of 1. ADHD session response rates are optimized above the diagonal, where positive prediction errors are more effective than negative signals. The impact of prediction error ratio $R_\omega = \omega_{\delta+}/\omega_{\delta-}$ on behaviour is illustrated in Figure 3.15. Similar to what was observed in the animal simulations, control group behaviour is best fit by $R_\omega \approx 1$, while ADHD behaviour is

optimized near $R_{\omega} \approx 1.4$. Furthermore, the mean response rates (Figure 3.9), IRTs (Figure 3.10), and regression slope (Figure 3.8, slope Table 3.2) emerging from the optimal "asymmetrical" model match those of the ADHD group reasonably well.

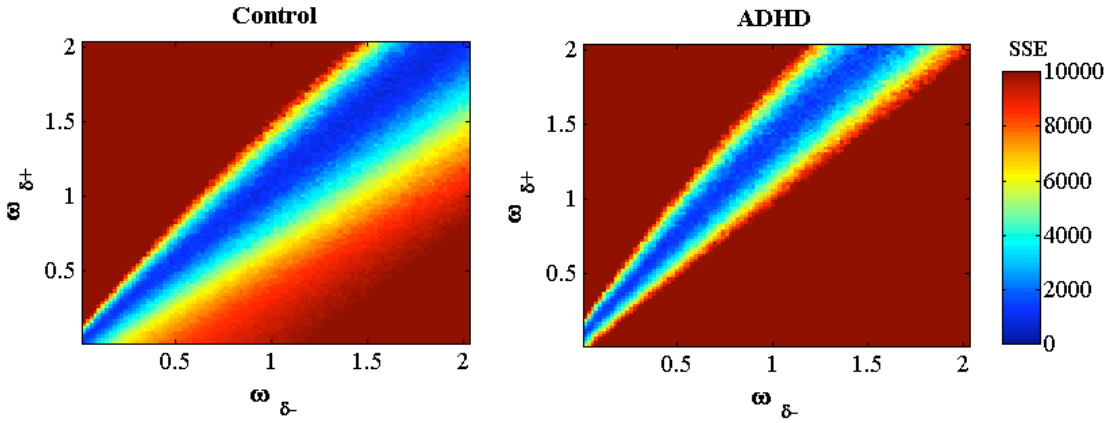

Figure 3.14: *Impact of positive and negative prediction error size*. Simulated behaviour compared to empirical control (left) and ADHD (right) group session response rates while varying the size of positive and negative prediction errors independently. *x-axis*: size of negative prediction error ($\omega_{\delta-}$). *y-axis*: size of positive prediction error ($\omega_{\delta+}$). *Colour*: Sum of squared error between simulated and empirical session response rates. Blue indicates lower error while red indicates higher error.

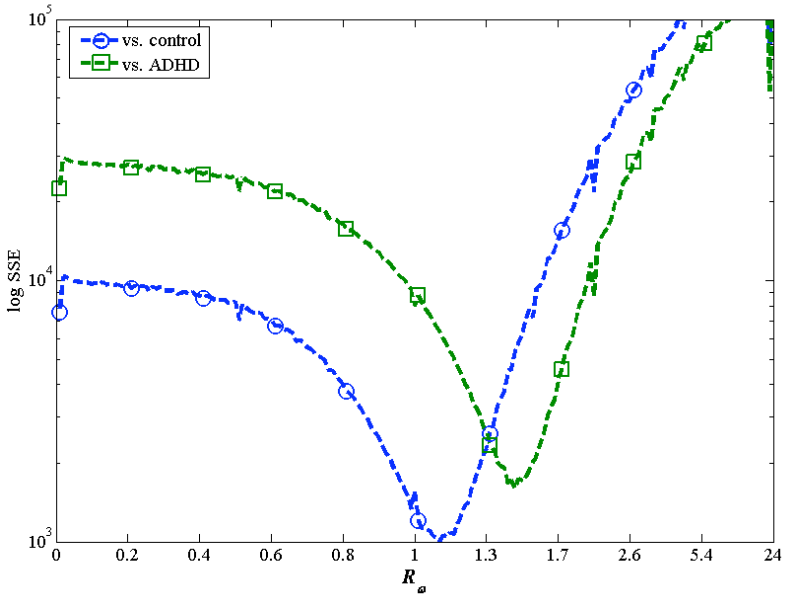

Figure 3.15: *Impact of prediction error ratio*. Simulated behaviour compared to empirical control (blue dashed line) and ADHD (green dashed line) group session response rates relative to the ratio of positive and negative prediction error size, $R_\omega = \omega_{\delta+} / \omega_{\delta-}$. Ratios were calculated for all values illustrated in Figure 3.14 ($0 \leq \omega_{\delta+}$, $\omega_{\delta-} \leq 2$), sorted, and pooled into 200 bins. *x-axis*: mean ratio of each bin. *y-axis*: Log sum of squared error between simulated session response rate averaged within each ratio bin, and empirical session response rates.

Finally, the search algorithm found a set of parameter values independent of prediction error size optimizing ADHD session response rates to define the "ADHD fit" model (SSE Table 3.2). An investigation into this solution reveals behaviour largely driven by exploration due to a high "temperature" value, τ, similar to what was found in the animal simulation (see Figure 3.11). Indeed, the mean response (Figure 3.9), and IRT (Figure 3.10) behaviour of the human "ADHD fit" model matches that of the ADHD group. However, session response rates expose a dissociation between human "ADHD fit" model and empirical behaviour (Figure 3.8). Session response rates of the ADHD group increased as a function of session during both fixed-interval and extinction components; whereas "ADHD fit" model response rates actually show a slight decrease (slopes Table 3.2).

A simulation of the human tasks outlined in Sagvolden et al. (1998), reveals two potential mechanisms driving elevated mean response rates, just as was observed in the animal simulations. However, an analysis of behavioural acquisition shows that only the "asymmetrical" model exhibits increasing response rates as observed in the empirical data. An investigation into the values being learned by the "asymmetrical" and "ADHD fit" models exposes the basis for this difference (Figure 3.16). The "ADHD fit" model has a high rate of exploration, meaning there is an increased probability of response even at states with low values. As such, it exhibits elevated response rates matching the behaviour of children with ADHD. However, the "ADHD fit" model has an intact learning mechanism; hence, it is capable of pushing earlier state values down in a "normal" direction in an attempt to compensate for an elevated degree of exploration (Figure 3.16). As a result, the "ADHD fit" model actually exhibits a slight decrease in response rate as a function of session (slopes Table 3.2). The "asymmetrical" model, on the other hand, has a compromised learning system. This model is incapable of sufficiently reducing state values due to a deficient negative prediction error signal (Figure 3.16). Hence, while the "ADHD fit" model responded excessively due to elevated exploration, the "asymmetrical" model learns ADHD-like behaviours due to elevating state values making a response increasingly likely as reinforcers are delivered.

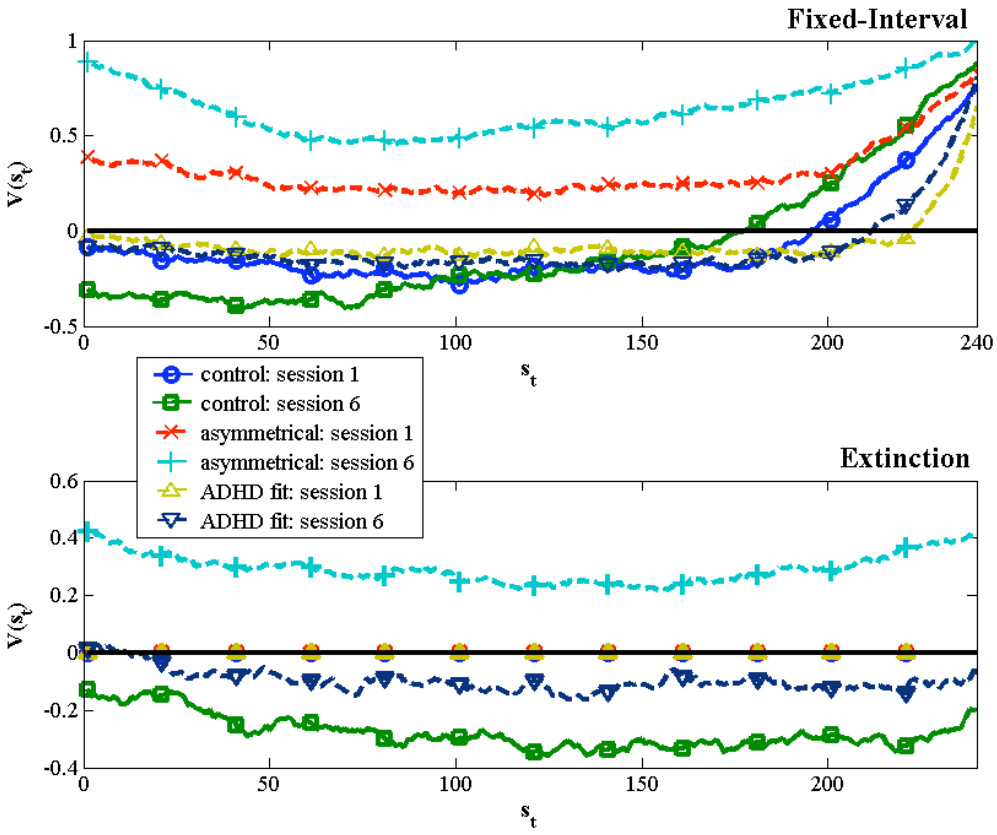

Figure 3.16: *Learning state values during the human multi-FI/EXT task simulation*. Fixed-interval (above) and extinction (below) state values after sessions 1 and 6 for "control", "asymmetrical" and "ADHD fit" models. Learning is demonstrated via the change in state values as the model experiences the environment. *x-axis*: all component states. *y-axis*: The value, $V(s_t)$ for each state early and late in learning.

To summarize, both the animal and human simulations show that a uniform prediction error scaling does not induce ADHD response behaviour. The animal simulation reveals two potential mechanisms that lead to increased stabilized hyperactivity, impulsivity and inattentive behaviours; namely, a system with asymmetrical prediction errors such that $R_\omega = \omega_{\delta+}/\omega_{\delta-} > 1$ or a system exhibiting high action exploration. The human simulation focusing on behavioural acquisition was able to demonstrate that only the asymmetrical prediction error model is able to account for the behavioural changes of children with ADHD.

4 Discussion

4.1 Principal findings:

Results from both animal and human multi-FI/EXT task simulations show that the prediction error's absolute size, as scaled by a factor $0 \leq \omega_\delta \leq 2$, does not play a significant role in bringing about ADHD-like behaviours (Figure 3.4 & Figure 3.12). Measures of the "hyper-/hypo" model's hyperactivity (FI response rate in Figure 3.1 & Figure 3.9), impulsivity (Figure 3.2 & Figure 3.10), and inattentiveness (EXT component response rate in Figure 3.1 & Figure 3.9) match control group behaviour far better than that of the ADHD group. Based on these simulation results, I conclude that neither abnormally large nor abnormally small prediction errors alone play a role in ADHD so long as both positive and negative error signals are balanced.

While these results provide an empirical argument against δ's size inducing ADHD-like behaviour in a multi-FI/EXT task, I provide theoretical evidence to this effect as well. A uniform scaling of both positive and negative prediction error signals is tantamount to adjusting the learning rate:

$$\begin{aligned} V\ (s_t) &= V\ (s_t) + \omega_\delta \eta \delta e(s_t) \\ &= V\ (s_t) + (\omega_\delta \eta) \delta e(s_t) \\ &= V\ (s_t) + \eta' \delta e(s_t) \end{aligned}$$

So long as $0 \leq \eta' < 1$, and the learning task is a Markov decision process (MDP), it has been proven that state values converge as the number of learning trials, $n \to \infty$ (Dayan & Sejnowski, 1994; Sutton, 1988). As I discussed previously, my simulations of the multi-FI/EXT task meets the criteria of an MDP. Furthermore, $\eta' = \omega_\delta \eta$ satisfies the criteria $0 \leq \eta' < 1$ for all model parameter settings, given $0 \leq \omega_\delta < 2$ (see Table 3.1 & Table 3.2). Therefore updating state values using either η, or $\eta' = \omega_\delta \eta$ will converge to the same

mean values. Altering the learning rate alone will not produce stabilized behavioural differences between control and ADHD groups reported by Sagvolden et al. (1992).

A set of parameters, independent of prediction error size, was found to produce ADHD-like mean response rates. The "ADHD fit" model exhibited mean response rates similar to that of animal and human ADHD groups (Figure 3.1 & Figure 3.9) due to high response exploration[3], facilitated by a high temperature value (Table 3.1 & Table 3.2). This emphasis on exploration over exploitation could also be labelled response variability since learned values have little impact on behaviour. Indeed response variability has been suggested as a core dysfunction in ADHD, lending support to a "temperature"-like dysfunction in ADHD (Russell et al., 2006), most likely rooted in noradrenaline dysfunction (Arnsten, Steere, & Hunt, 1996; M. J. Frank, Santamaria, O'Reilly, & Willcutt, 2006). However, the "ADHD fit" model's response rate across session did not match that of children with ADHD (Table 3.2 & Figure 3.8). Since the "ADHD fit" model's learning mechanism is intact it is able to push values for non-rewarding states down to "normal" levels (Figure 3.16), resulting in behavioural changes similar to those of controls (see slopes Table 3.2, Figure 3.8). Hence, exploration alone cannot account for the increasing hyperactivity, impulsivity and inattentiveness of children with ADHD. The behavioural changes observed in the ADHD group appear systematic, which suggests some abnormality in the learning mechanism itself.

In line with this, simulation results reveal a possible link between abnormal prediction errors and ADHD in the form of an asymmetry between positive and negative error signals. As illustrated in Figure 3.7 and Figure 3.15, control group behaviour is fit best when positive and negative prediction errors are balanced such that $R_{\omega} = \omega_{\delta+}/\omega_{\delta-} \approx 1$. Balanced positive and negative error signals, $\omega_{\delta+} \approx \omega_{\delta-}$, ensure that state values reflect favourable and unfavourable outcomes equally. Conversely, ADHD group behaviour is fit best when there is an asymmetry between positive and negative error signals, such that

[3] The animal "ADHD fit" model also has a higher decay rate value than the control model (Table 3.1). However, decay rate, along with learning rate, was demonstrated to play an minor role in response behaviour (Figure 3.3). The decay rate value set for the "ADHD fit" model was one of many possible values that would provide an equally good fit.

$\omega_{\delta+} > \omega_{\delta-}$. The ratio of prediction error scaling factors, $R_\omega = \omega_{\delta+}/\omega_{\delta-} > 1$ was demonstrated to provide the best fit in terms of stabilized behaviour (Figure 3.7) and the rate of behavioural change (Figure 3.15). Asymmetrical signals instil a bias in the direction of the asymmetry; hence, a system with $\omega_{\delta+} > \omega_{\delta-}$ will learn state values that overemphasize positive reinforcement. This bias will be exhibited by behaviour that appears ignorant of negative environmental factors relative to the positive reinforcement encountered. This is precisely what is observed in the "asymmetrical" model's behaviour.

4.2 Functional level description

In a normal system, where $\omega_{\delta+} \approx \omega_{\delta-}$, the positive value of reinforcers delivered at the end of each fixed-interval trial is tempered by the response cost. This is not the case in a system with $\omega_{\delta+} > \omega_{\delta-}$. Here, positive reinforcer value is absorbed excessively by previously encountered state values, which cannot be effectively reduced by the response cost factor (Figure 3.16). The probability of response increases in correspondence with increasing state values. Thus, hyperactive behaviour emerges, measured as mean response rate during fixed-interval components. Furthermore, increasing the number of responses during a fixed time interval necessarily reduces the time delay between responses. Hence, IRTs are reduced, increasing the motor impulsivity measure. Like hyperactivity, attention deficit, defined as excessive responding during the extinction component, develops due to increasing values in extinction component states.

The mechanism driving increased attention deficit measures in my model is not the same at the mechanism proposed by the hypo-active theory of dopamine in ADHD. The hypo-active dopamine theory hypothesized that dopaminergic dysfunction in ADHD gives rise to a shorter and steeper delay-of-reinforcement gradient. This, they argue, prevents an appropriate binding of the context, behaviour and reinforcer (Figure 1.2) (Sagvolden et al., 2005). Hence, they propose that children with ADHD are more likely to exhibit behaviour that has been rewarded previously, independent of the current context. First, I note that the specific mechanisms involved in the delay-of-reinforcement

gradient dysfunction are unclear. Factors such as learning rate, discount factor and eligibility trace strength are confounded into a single mechanism of reinforcer impact on behaviour. However, figures Figure 3.3 and Figure 3.11 show that varying any these three factors does not result in ADHD-like behaviour.

The "asymmetrical" model, on the other hand, develops "attention deficit" for different reasons entirely. As noted earlier, an excessive proportion of reinforcer value is propagated back through fixed-interval states when $\omega_{\delta+} > \omega_{\delta-}$. As a result, states near the start of the fixed-interval state space are endowed with inappropriately high values (Figure 3.16). The value of the first fixed-interval state is then propagated back through extinction states via transition T_3 (Figure 2.4). Hence, the "asymmetrical" model not only propagates reinforcer value back through fixed-interval states inappropriately, but through extinction states as well. As such, the model develops "attention deficit" by inappropriately crediting extinction states with reinforcers awarded during the fixed-interval component. In the context of the multi-FI/EXT experiment, it is as though the "asymmetrical" model has inappropriately associated response states during the extinction component with the start of the fixed-interval component because the extinction component predicts the fixed-interval component. Hence, "attention deficit" is not properly due to an attentional dysfunction; rather, it emerges from exaggerated state values caused by a deficient negative prediction error that cannot retard the propagation of positive reinforcement.

Finally, note that both control and ADHD optimal fits straddle the "normal" prediction error size, where $\omega_{\delta+}=\omega_{\delta-}=1$, across a range of scaling factors along the diagonal (Figure 3.6 & Figure 3.14). This suggests that the size of the prediction error, and hence the learning rate, does not play a crucial role in defining response behaviour on a multi-FI/EXT task. Though negligibly small learning rates, $\eta \approx 0$, do not provide a good fit to the empirical data since the model is incapable of learning within the limited number of trials, $\eta > 0.1$ is capable of matching mean response rate and response rate changes across sessions (Figure 3.3 & Figure 3.11). While this offers evidence that ADHD may be associated with a dopamine signal asymmetry independent of absolute signal size, it

also suggest that the multi-FI/EXT task is relatively insensitive to learning rate. Hence, it is possible that ADHD is associated with abnormalities in both relative and absolute dopamine signal efficacy. Though it is difficult to isolate absolute learning rate from a learning rate asymmetry, I propose that the rising optimum task may provide some insight into learning rate abnormalities associated with ADHD as it can be used to provide a measure of information integration over variable trial lengths (Herrnstein, 1991). Should children with ADHD have abnormally large prediction errors their response behaviour will depend on recent trial outcomes, whereas small prediction errors will facilitate behaviour that considers a longer trial history. Indeed, investigations have revealed that children with ADHD are influenced more by recent reinforcement than their previous reward history (Tripp & Alsop, 1999), suggesting an elevated learning rate, which may be the result of abnormally large prediction errors.

4.3 Biological implications:

The previous analyses and discussion have focused on a theoretical association between asymmetrical prediction errors and ADHD. However, the question remains: What are the neurological mechanisms through which this asymmetry might emerge, and how might this asymmetry alter the function of the various neurological systems influenced by dopamine?

The model discussed in this thesis abstracts large swaths of biological detail. While a high level of abstraction allows for a broad investigation of dopamine ’s role in ADHD, it must also be interpreted with caution. A closer inspection shows that there are multiple points at which signal asymmetry could be introduced into the system: 1) signal generation in the midbrain dopamine system, 2) during communication between dopamine neurons and their targets, and 3) at the target neuron where neurological changes occur as a result of dopaminergic input. Furthermore, due to the intimate relationship between learned values and response selection, dysfunction at one or more of these points could induce not only reinforcement learning deficits, but action selection

deficits as well. In the following I address some of the biological issues my modelling results highlight.

4.3.1 Generating asymmetrical prediction errors

The hyper-active dopamine theory of ADHD proposed by Grace (2001) offers evidence associating ADHD with a dysfunction at the point of prediction error generation in the midbrain dopamine system. Here, Grace suggests that reduced afferent excitation leads to reduced self-inhibition of the midbrain dopamine system. This culminates in a hyper-active midbrain dopamine system producing abnormally large prediction errors. However, this hypothesis focuses primarily on positive prediction errors, leaving negative prediction errors largely unaddressed. Grace (2001) suggests that negative prediction errors will be exaggerated in correspondence with the exaggerated positive values being learned due to abnormally large positive prediction errors. However, it is not at all clear that this should be the case. This treatment of negative prediction errors overlooks the base assumption that dopamine neurons are hyper-active due to low concentrations of extra-cellular dopamine. Given their hyper-active state, dopamine neurons may have difficulty halting their activity to encode negative prediction errors properly. Indeed, it seems reasonable to suggest that negative prediction errors would be blunted in correspondence to the degree that the dopamine system is hyper-active. This provides some evidence that an asymmetrical effect of prediction errors may be rooted in the midbrain dopamine system itself.

4.3.2 DAT's role in asymmetrical striatal learning

Dopamine transporter (DAT) is the primary mechanism through which dopamine reuptake occurs in striatum. Given that reuptake removes the dopamine signal from the system, DAT is a central factor modulating signal transmission,. ADHD is associated with an over-expression of DAT in striatum (Krause et al., 2003; Madras, Miller, & Fischman, 2002). This offers a mechanism through which extra-cellular concentrations of

dopamine may be reduced, extending or providing an alternative to the frontal afferent hypothesis forwarded in the hyper-active dopamine theory of ADHD.

Of particular interest, recent modelling work has demonstrated that DAT is capable of limiting the sphere and effective lifetime of extra-cellular dopamine (Cragg & Rice, 2004). Given the evidence that dopamine acts to increase the signal-to-noise ratio by allowing only the strongest activity to persevere (Cohen & Servan-Schreiber, 1992; Hernandez-Lopez, Bargas, Surmeier, Reyes, & Galarraga, 1997), a DAT induced decrease in dopaminergic influence would likely increase striatal noise. Furthermore, it was demonstrated that striatal neurons expressing high-affinity dopamine receptors (e.g. inhibitory D2 receptors) would be exceptionally sensitive to DAT induced variation in extra-cellular concentrations of dopamine (Cragg & Rice, 2004). This suggests that populations of D2-expressing striatal neurons will exhibit disproportionately noisy activity in the face of reduced extra-cellular dopamine.

In order for a reinforcement learning signal to be effective it requires some degree of stability in what is being learned. Reinforcement learning performance will degrade as variance in either state identity or state value increases. More formally, a reinforcement learning algorithm cannot be guarantied to converge if the state signal is non-Markovian (Sutton & Barto, 1998). While a state signal is formally either Markovian or it is not, it is useful to think of Markov property violations along a continuum from slight to strong. Given this continuum, one will obtain better learning performance to the extent that the Markov property holds (Sutton & Barto, 1998).

I propose that noise in a neural reinforcement learning system can be viewed as state signal contamination. Hence, as noise increases the Markov property is increasingly violated, leading to impaired learning performance. Assuming this, it follows that a DAT induced increase in striatal noise specific to populations of D2-expressing neurons would reduce that substructure's ability to learn effectively from prediction errors encoded in the phasic activity of the midbrain dopamine system. Should this be the case, D1-expressing striatal neurons would be more adaptive relative to their D2-expressing

counterparts. In line with the asymmetrical prediction error efficacy suggested by my model, low-affinity D1-expressing striatal neurons are thought to be more responsive to phasic bursts of dopamine relative to phasic dips (M. J. Frank, 2005; Richfield, Penney, & Young, 1989). Hence, a DAT induced asymmetry could result from increased learning in D1-expressing neurons relative to D2-expressing neurons.

4.3.3 Dopamine and response selection

The hyper-active dopamine theory of ADHD provides evidence that reduced afferent excitation of striatal neurons plays a prominent role in ADHD (Grace, 2001). This theory focuses on the impact such a dysfunction would have on extra-cellular dopamine concentrations. However, frontal stimulation also modulates activity in striatal neurons themselves (Wilson & Kawaguchi, 1996).

Medium spiny neurons (MSNs) in the striatum are thought to be bi-stable, existing in either an up- or down-state of activity (Wilson & Kawaguchi, 1996). Up-state neurons are characterized by a relatively high spike-rate while down-state neurons exhibit reduced, sporadic spike-rates. Whether a particular neuron is in an up- or down-state has been shown to be dependent on convergent excitation from frontal systems (O'Donnell & Grace, 1995; Wilson & Kawaguchi, 1996). As afferent excitation increases, a neuron will move into an up-state, and conversely, a decrease in convergent stimulation will push the neuron into a down-state. This acts to gate information through the striatum, where striatal output will predominantly reflect the activity of neurons receiving strong convergent information from afferent sources (Grace, 2001). Given that the striatum has been shown to play a major role in action selection, information gating can loosely be thought of a selecting one action among multiple competing options (M. J. Frank, 2005; Mink, 1996).

Dopamine is thought to interact with MSN state such that D1-expressing down-state neurons will be inhibited in the presence of dopamine, while D1-expressing up-state

neuron activity will increase (Hernandez-Lopez et al., 1997). Through this mechanism dopamine is said to enhance the contrast between up- and down-state D1-expressing neurons, increasing the signal-to-noise ratio of striatal neurons.

Relating these dynamics back to ADHD, reduced frontal activity has been associated with the disorder (Grace, 2001; Sieg, Gaffney, Preston, & Hellings, 1995; Zametkin, Nordahl, Gross, & King, 1990; Zametkin et al., 1993). This suggests that the proportion of up-state neurons may also be reduced, increasing the proportion of activity in striatal output produced by down-state neurons. Reduced concentrations of extra-cellular dopamine will exacerbate the noise in striatal output as the contrast between up- and down-state D1-expressing MSNs will be compromised. As previously discussed, increased striatal noise could reduce the systems ability to learn. Additionally, I propose that reduced contrast between up-/down-state D1-expressing MSNs will result in sub-optimal action selection. This could produce impulsive behaviour as observed in children with ADHD should actions be selected at inappropriate times, or inappropriate actions be selected for the given situation. Hence, excessive contribution from down-state neurons in striatal output, initially driven by insufficient afferent striatal stimulation, could provide a mechanism for the hyperactive and impulsive behaviours associated with ADHD.

4.3.4 Salience and discounting

Finally, building upon ideas related to reduced frontal activity in ADHD, recent imaging work shows that both frontal and striatal systems are involved when making decisions related to delayed rewards (McClure, Laibson, Loewenstein, & Cohen, 2004). The relative activity of what can loosely be considered frontal and striatal systems was shown to predict subjects' decisions. Specifically, subjects choose small immediate rewards when the striatal system was more active than the frontal system, and choose larger delayed rewards when the opposite was true. Hence, children with ADHD may

excessively discount the value of future rewards due to a hypo-active frontal circuit responsible for calculating the value of delayed rewards.

Some evidence suggests that indeed, children with ADHD do discount future rewards excessively. Children diagnosed with ADHD were shown to have abnormal electrophysiological responses to rewards (C. B. Holroyd, Baker, Kerns, & Müller, 2008). It was proposed that reward salience plays a key role in that abstract rewards (points) did not elicit typical electrophysiological responses; however, when real rewards (money) were provided the electrophysiological responses began to normalize. This raises the possibility that children with ADHD discount future rewards excessively, observed in this study as monetary rewards predicted by points. In addition, a study involving young adults showed temporal discounting difference between ADHD subjects and controls, but only when real rewards were provided (Scheres, Lee, & Sumiya, 2008). Again, this suggests reward salience and abnormal discounting may play a key role in the impulsivity associated with ADHD, an idea long thought to be central to the disorder (see Barkley (1997a)). These results indicate that some of the symptoms associated with ADHD could be the result of an atypical interaction between striatal and frontal circuits involved in determining action utility. Exploring the details of this interaction between frontal and striatal systems, and how this may couple with the model presented in this thesis present an interesting future direction of investigation.

In summary, I have outlined several potential biological mechanisms that could give rise to my model's prediction of an asymmetrical effect of dopamine. As discussed above, such a dysfunction could arise from asymmetrical signal generation, or due to a reduction in the signal-to-noise ratio in striatum. Atypical activity in these systems may impact both learning and action selection. Furthermore, evidence suggests that the relative activity of frontal and striatal circuits plays a central role in action utility, which may be compromised in ADHD. Clearly each of these mechanisms in isolation is complex, and any potential interactions among them even more so. Hence, a more detailed empirical and model-based investigation of their potential role in ADHD is required, but beyond the scope of the work presented in this thesis.

4.3.5 Effects of medication

The primary drug used to treat ADHD, methylphenidate (e.g. Ritalin), acts via DAT blockade. By reducing the efficiency with which dopamine is removed from the system, it is thought that system function normalizes as extra-cellular levels of dopamine increase (Grace, 2001). Importantly, these drugs have both immediate and learning-related effects; suggesting mechanisms beyond abnormal reinforcement learning signals are associated with ADHD. While a formal treatment of the impact and dynamics of dopaminergic drug treatment is not possible here, I suggest that the mechanisms through which drug treatment is thought to operate could normalize asymmetrical prediction error efficacy through various methods.

As discussed previously, one potential source of prediction error asymmetry stems from the hyper-active midbrain dopamine system's inability to cease firing and properly encode negative prediction errors. Drugs that increase extra-cellular dopamine will inhibit the midbrain dopamine system (Grace, 2001). Increased inhibition would facilitate the dopamine system's ability to cease its activity to properly encode negative prediction errors, normalizing an asymmetrical encoding of prediction errors.

DAT blockade should also normalize noise levels in striatum, which as discussed previously, could occur via an over-expression of DAT or an interaction between MSN state and dopamine. Blocking DAT reduces the proportion of effective transporter membranes, which would regulate extra-cellular dopamine's impact on D2-expressing MSNs. This would allow the prediction error signal to operate on this sub-system more effectively as activity in that population of neurons would be more stable and representative of optimal action selection. Furthermore, increasing the concentration of extra-cellular dopamine will also facilitate increased contrast enhancement in D1-expressing MSNs. This should act to improve the reliability of response gaiting which may reduce impulsive and hyperactive behaviours. As improvements in action selection

are not dependent upon learning, this provides a potential explanation for the immediate effects of drugs like methylphenidate, which cannot be due to changes in learning.

In closing this discussion on the biological implications of my model, I caution that the relationships between extra-cellular concentrations of dopamine, tonic and phasic dopamine activity are very complex. Before any concrete prediction can be made, a more formal and model-based approach explicitly examining these issues, along with the impact of DAT blockade are required. However, evidence presented here suggests several plausible biological mechanisms through which an asymmetrical effect of prediction error signals might emerge.

4.4 Comparison with other models

Two computational models focusing on dopamine system dysfunction in ADHD have proposed prediction error abnormalities that are not in line with my findings discussed in this thesis. Based on a TDRL application to a delayed response time task, Williams and Dayan (2005) focus on an association between ADHD and an abnormal learning rate. Work by Frank et al. (2006) advances the hypothesis that ADHD is associated with ineffective positive prediction errors. I address these two models in the following.

4.4.1 Impulsivity in a delayed response time task

Sonuga-Barke, Taylor, Sembi, & Smith. (1992), have suggested that children with ADHD are delay averse, which presents itself behaviourally as impulsivity. Children took part in a delayed response time task in which they were asked to choose either a small immediate reward or a large delayed reward. In the *post delay condition*, the total delay associated with both choices was made the same by including a post reward delivery delay. A long delay followed small immediate rewards, and a short delay followed large delayed rewards such that the total delay for both reward choices was equal. In this condition, both controls and children with ADHD performed optimally by choosing the

large delayed reward more often. This demonstrated that children with ADHD are capable of maximizing reward just as well as controls in the face of an impulsive option. However, in a *trials constraint condition* the post reward delay was removed and the number of potential rewards was limited. In this condition children with ADHD performed sub-optimally by choosing the small immediate reward, whereas controls preferentially selected the large delayed reward. Hence, it was concluded that children with ADHD are delay averse, choosing to optimize the time between reward deliveries rather than the total amount of reward.

Williams and Dayan (2005) applied a TDRL model to a simulation of the trials constraint condition of the delayed response time task. In doing so, they explored the role of learning rate, discount rate, action bias[4], and temperature parameters. They concluded that the impulsive behaviour reported by Sonuga-Barke et al. (1992), can be caused by an abnormalities in learning mechanisms (learning rate), in use of learned information (action bias/temperature), or information available to be learned (discount rate). Of particular relevance, they replicated ADHD group performance by reducing the learning rate by a factor of 50%, effectively modelling stunted phasic dopamine activity as suggested by the hypo-active dopamine theory of ADHD (Sagvolden et al., 2005). The modelled "control" group increasingly selected the large delayed reward, choosing the large reward 60% of the time by the end of the simulation as observed in the empirical data. The modelled "ADHD" group, on the other hand, chose the delayed reward only 20% of the time. They demonstrated the when the learning rate is too low, the model cannot reliably learn about the delayed reward's value given the limited number of trials. Setting the learning rate too high results in over-learning the immediate reward's value, significantly reducing the probability that it will ever encounter the large delayed reward. To investigate the impact of learning rate on stabilized behaviour, the experiment was extended by increasing the number of trials from 30 up to 150. Results on the augmented task showed that ADHD behaviour, modelled as a reduced learning rate, should

[4] Action bias amounts to the model's initial preference for action vs. inaction. This is very similar to the response threshold parameter, $\phi = 1$, used in my model. Though this parameter is locked in my model, state values are learned in reference to it; hence, the initial value is of little importance.

eventually asymptote at normal performance levels where the optimal delayed reward is selected most of the time.

This prediction is in line with the previous convergence discussion showing that as $n \rightarrow \infty$, a system with an altered learning rate, $0 \leq \eta' < 1$, will converge to the same asymptote as any other learning rate within the same bounds. However, this prediction is contrary to the stabilized response behaviour of the animal ADHD-model group in Sagvolden et al. (1992). In my simulations, I demonstrate that acquisition and stabilization of ADHD group behaviour on the multi-FI/EXT task cannot be induced by atypical learning rate, discount factor, or temperature values as suggested by Williams & Dayan (2005). I propose that only a prediction error ratio, $R_\omega = \omega_{\delta+}/\omega_{\delta-} > 1$, drives performance similar to that of ADHD groups on this task. I further suggest that this finding can be extended to explain sub-optimal performance on the delayed response time task as observed in children with ADHD.

In the trials constraint conditions, subjects must learn to wait for the larger reward, requiring them to pass over the temptation of an immediate reward. The control group was shown to be capable of doing so after some experience with the task (Sonuga-Barke et al., 1992). Within a TDRL framework this implies that the value of the large delayed reward has had the opportunity to propagate from the time of reward delivery back to the start of the trial where the choice is made, as demonstrated in (Williams & Dayan, 2005) (see Figure 2). With each experience of the large delayed reward, preceding state values increase above the predicted value of the small reward via positive δ's. Hence, selecting the small immediate reward will result in a reinforcer smaller than the predicted value (the large reward), resulting in a negative δ for that choice. A "normal" model with $R_\omega \approx 1$, can use positive and negative δ's equally to learn optimal performance. However, if the system is geared such that positive prediction errors are more effective that negative signals, the model will have difficulty devaluing the immediate reward in relation to the delayed reward. Such a system will only be able to learn from the positive δ's associated with selecting the large delayed reward, not the negative δ's that result from selected

small immediate rewards. As a result, the "asymmetrical" model will be slower to achieve optimal performance rates, if it can achieve them at all.

I conclude that an asymmetrical prediction error, $R_{\omega} = \omega_{\delta+}/\omega_{\delta-} > 1$, can not only replicate ADHD behaviour on a multi-FI/EXT task, but that it will be capable of replicating ADHD behaviour on the trials constraint condition of the delayed response time task. However, future work should explore the impact of an asymmetrical system on all conditions of this task to show that "control" and "ADHD" models match performance on other task conditions as well.

4.4.2 Learning from positive and negative feedback

A dopaminergic explanation of ADHD has been investigated using a computational model simulating dynamic dopamine interactions within circuits linking the basal ganglia and frontal cortex (M. J. Frank et al., 2006). This model has been used to provide insights into Parkinson's disease (M. J. Frank, 2005), orbitofrontal damage (M. J. Frank & Claus, 2006), in addition to numerous prediction regarding the nature of dopaminergic dynamics of systems interacting with the basal ganglia (M. J. Frank, Woroch, & Curran, 2005; M. J. Frank & O'Reilly, 2006; M. J. Frank, Seeberger, & O'Reilly, 2004).

This model learns to choose responses that lead to positive reinforcement while suppressing responses that lead to negative outcomes. Each possible response has corresponding "Go" and "NoGo" representations encoded by unique pathways of activation. Based on the relative activation of these pathways, a response will either be chosen ("Go") or avoided ("NoGo"). The strength of "Go"/"NoGo" representations are learned using dopamine signals. Dopamine input to "Go"/"NoGo" units is normally held at a baseline activity level of 0.5, representing tonic stimulation. After a correct response dopamine levels are increased from baseline to 1.0, simulating a phasic burst. This burst strengthens "Go" pathway activity while simultaneously suppressing the "NoGo" pathway activation, making a "Go" response more likely in the future. Conversely, an

incorrect response is met with reduced dopamine input, dropping below baseline to 0.0, simulating a phasic dip. This strengthens the "NoGo" pathway and inhibits the "Go" pathway, making that response less likely in the future. Hence, dopamine modulation facilitates reciprocal "Go"/"NoGo" learning to maximize positive reinforcement. Furthermore, tonic dopamine activity plays a critical role in defining the dynamic range within which phasic signals can operate, and providing a constant biasing signal onto striatal units.

A model-based investigation of dynamic dopamine activity shows that normal levels of tonic dopamine are required for balanced positive and negative reinforcement learning, which is disrupted in Parkinson's disease (M. J. Frank, 2005). Parkinson's disease is cause by the death of dopamine neurons, leading to a dramatic reduction in tonic dopamine levels. It was demonstrated that a reduction in tonic dopamine, as in the case of Parkinson's disease, leads to improved learning from negative feedback and impaired learning from positive feedback (M. J. Frank et al., 2004). Furthermore, tonic levels above baseline, as caused by medication, lead to improved learning from positive feedback and impaired learning from negative feedback. Hence, tonic dopamine levels were demonstrated to play a critical role in biasing learning from positive and negative outcomes.

Using the "Go/NoGo" model of the basal ganglia, ADHD was modelled by lowering tonic dopamine below "normal" levels of activity (M. J. Frank et al., 2006), precisely the manipulation used to model Parkinson's disease (M. J. Frank, 2005). This leads to chronic "NoGo" pathway excitation and corresponding "Go" pathway inhibition. Although phasic dopamine bursts elicit a "normal" absolute positive difference above tonic activity levels, "Go" learning is impaired as positive bursts are unable to counteract the consistent inhibitory bias driven by low tonic stimulation. Conversely, due to a floor effect, the range below abnormally low tonic activity level is reduced. However, due to tonically induced "NoGo" learning bias, "NoGo" learning is improved despite a reduction in the absolute size of phasic dips. Hence, this model of ADHD predicts a bias favouring "NoGo" learning and impairing "Go" learning since positive dopamine bursts

will be unable to sufficiently drive “Go” activation in the face of consistent “NoGo” pathway stimulation. Furthermore, the model suggests that “Go” learning should be improved by drugs that remove the “NoGo” activation bias by elevating tonic dopamine levels, such as methylphenidate.

The same probabilistic selection task used to investigate Parkinson’s disease in (M. J. Frank, 2005) was applied in investigation of ADHD (M. J. Frank et al., 2006). In each trial, the subject is forced to choose one of two symbols, where each symbol is associated with a different probability of reward. During the training phase, three pairs of stimuli are learned, each with unique probabilities of reward: A/B (80/20%), C/D (70/30%), and E/F (60/40%). After each response, feedback is provided noting the subject’s choice as “correct” or “incorrect” based on the reward contingencies of the selected stimuli. Once performance meets specified performance criteria the test phase begins. During testing, novel stimuli pairs are presented (e.g. A/C, or A/F) and no feedback is provided. A subject will choose stimuli with a high probability of reward (e.g. ‘A’ or ‘C’) to the degree that they learned from positive feedback. Conversely, a subject will avoid stimuli with a low probability of reward (e.g. ‘B’ or ‘D’) to the degree that they learned from negative feedback. This, it is argued, provides a behavioural measure of the relative efficacy of positive and negative dopamine signals on “Go” and “NoGo” pathways in the basal ganglia. Hence, “choose A” performance provides a measure of “Go” pathway learning, whereas “avoid B” performance provides a measure of “NoGo” pathway learning.

Empirical findings on the probabilistic selection task show that adult subjects with ADHD while off medication (ADHD OFF) are impaired in both “choose A” and “avoid B” conditions (Figure 2, (M. J. Frank et al., 2006)). Subjects on medication (ADHD ON) showed significant “choose A” performance gains, but no change in “avoid B” performance. These results are interpreted as demonstrating that a drug-induced increase in tonic dopamine acts to improve “Go” learning via increased dopamine burst efficacy, exhibited by the selective “choose A” improvements in the ADHD ON group. This is

taken further to suggest that ADHD is associated with impaired "Go" learning caused by ineffective positive prediction errors, which medication acts to relieve.

This interpretation is a sharp contrast to my interpretation of the multi-FI/EXT simulation results. I argue that ADHD is associated with a prediction error asymmetry such that $\omega_{\delta+} > \omega_{\delta-}$. This is the polar opposite of the proposed dysfunction reported by M. J. Frank et al. (2006), which suggests an association between ADHD and a prediction error asymmetry such that $\omega_{\delta+} < \omega_{\delta-}$. While the explanatory power of the BG model is convincing, there are several factors that make a comparison between these results and my own difficult.

Subjects in the Frank et al. (2006) study were adults taken off medication 24-hours prior to the experiment. The behaviour used to constrain my model was collected from medication naïve children. This introduces factors such as age, long-term effects of medication and effects of being taken off medication that differ between this study and the experiments reported in Sagvolden et al. (1998).

In addition, the BG model predicts several results that aren't found in the empirical data. The model predicts that low tonic dopamine (ADHD OFF) will be associated with improved "NoGo" learning and impaired "Go" learning. However, ADHD OFF subjects were shown to be equally bad at both positive and negative learning. The BG model also predicts that increased tonic levels of dopamine should be associated with improved "Go" learning and impaired "NoGo". While improved "Go" learning was observed in ADHD ON subject, "NoGo" impairment was not observed. This suggests that factors beyond tonic dopamine activity is involved in ADHD (which were accounted for in this model by including a norepinephrine dysfunction). A crucial factor that was not explored in this model of ADHD was the coupling between tonic and phasic dopamine activity. Though tonic levels of dopamine were modulated, corresponding effects on phasic activity were not included. This could be explored by including abnormally large and small phasic bursts of dopamine in correspondence with abnormally low and high tonic activity as suggested by Grace (2001). Furthermore, it should be noted that extra-cellular

concentrations of dopamine are not necessarily linearly related to tonic levels of dopamine activity at the synapse. Given that ADHD is associated with atypical DAT expression, it is possible that there is an atypical relationship between intra-synaptic and extra-synaptic tonic dopamine levels. Hence, it is not clear that reducing tonic activity is the best method to investigate the association between ADHD and low concentrations of extra-cellular dopamine.

Finally, modelling ADHD by lowering tonic levels of dopamine is precisely the method used to model Parkinson's (M. J. Frank et al., 2004). However, the symptoms associated with Parkinson's differ drastically from symptoms associated with ADHD, as do the proposed sources of dysfunction. While Parkinson's disease is caused by dopamine cell death, the midbrain dopamine structure is thought to be intact in ADHD, suggesting very different sources of dysfunction. In light of the previous discussion on the biological mechanisms that may give rise to an asymmetrical effect of prediction errors in ADHD, I would argue that additional parameters such as modulating the proportion of up-/down-state activity or introducing a DAT parameter may help the BG model better predict behaviour of subjects with ADHD. These additional parameters may help to explain why ADHD and Parkinson's produce vastly different symptoms despite both disorders being associated with low concentrations of extra-cellular dopamine.

4.4.3 Summary

In summary, I provide evidence suggesting that neither the model proposed by Williams & Dayan (2005), nor the model proposed by M. J. Frank et al. (2006) can account for the range of behavioural data discussed in this thesis. The learning rate model of ADHD, (Williams & Dayan, 2005), is incapable of developing atypical stabilized behaviour as reported by Sagvolden et al. (1992). The low tonic dopamine model presented by M. J. Frank et al. (2006) cannot explain "avoid B" performance deficits in the probabilistic selection task, nor can it clearly explain the impact of medication. I propose that asymmetrical phasic activity, where $\omega_{\delta+} > \omega_{\delta-}$, can reproduce the behaviours

of subjects with ADHD on the delayed response time as reported by Sonuga-Barke et al. (1992). While it is difficult to make any clear predictions about adult performance on the probabilistic learning task, my model would predict that children with ADHD naïve to medication would show benefits in “Go” learning and impaired “NoGo” learning in correspondence with more effective positive prediction errors. I should caution that I cannot make a strong prediction here given the complexity of the biological mechanisms involved in reinforcement learning. Future work should investigate these predictions more thoroughly.

5 Conclusions

The issues involved in ADHD are complex, both at a neurological and a behavioural level. While the results presented here provide a promising basis from which to explore related issues further, significant work remains before an understanding of the neurological mechanisms that give rise to the behavioural markers of ADHD can be achieved. Here, I present promising channels of future research that emerge from the results presented in this thesis, and concluding remarks.

5.1.1 Future research

Throughout this thesis I have hinted at several points from which this initial work investigating dopamine's role in ADHD may be extended. In particular, the model presented here could benefit from additional empirical constraints. Notably, asymmetrical prediction errors could be applied to the discounting tasks presented by Sonuga-Barke et al. (1992) and modelled by Williams & Dayan (2005), as well as the probabilistic selection task described by M. J. Frank (2005) and modelled by M. J. Frank et al. (2006). This would provide an explicit treatment of decisions involving delayed versus immediate rewards, as well as positive versus negative learning, neither of which are explicitly investigated in the multi-FI/EXT task discussed in this thesis.

The model presented here could be extended with two particularly relevant additions. As mentioned previously, two systems, one impulsive and the other premeditative, are thought to be involved when making decisions involving temporally delayed rewards (McClure et al., 2004). Furthermore, these two systems, striatal and frontal, are thought to be involved in ADHD dysfunction (Barkley, 1997a). My model collapses these two systems into a single decision making unit; however, the interplay between these two systems and their involvement on tasks such as the discounting task just discussed could be particularly enlightening.

In addition, my model assumes that the state space being operated on by the learning agent is the same for controls and ADHD. However, recent modelling work shows that this assumption may not be correct. Negative prediction errors have been suggested to play a role in defining the state space operated on by a learning agent (A. D. Redish, Jensen, Johnson, & Kurth-Nelson, 2007). This additional role attributed to negative prediction errors was shown to provide an account of rapid reinstatement of extinguished behaviour, a phenomena conspicuously unaccounted for by traditional TDRL techniques, as well as impulsive gambling and addiction relapse. Furthermore, a recent proposal argues that ADHD is the result of an inefficient value transfer from states where actual rewards are received back to states that predict reward (Tripp & Wickens, 2008). This theory does not focus on state-space creation whatsoever; however, it is possible that inefficient value transfer could be related to inefficient state-splitting driven by ineffective negative prediction errors. Varying the state-space between controls and ADHD could prove to be particularly informative, especially in terms of treatment and coping with the disorder.

Finally, the William and Dayan model of ADHD discussed by Williams & Dayan (2005) has been extended to include variables associated with state regulation and cognitive function (Williams & Taylor, 2004), and applied to a picture matching task. In this task subjects were asked which of many possible images matched a target image under variable time constraints (Sonuga-Barke, 2002a). The behaviour of children with ADHD was shown to match that of controls when the time constraint was limited to 5- or 15-seconds; however, ADHD subjects were notably impaired at the 10-second constraint condition. Williams and Taylor suggest that increased dopamine appetite, which is modelled as an increase in response variability whenever dopamine levels are outside a tightly bounded “optimal” level, produce ADHD-like behaviour on the matching task (Williams & Taylor, 2004). In addition they suggest that cognitive deficits, modelled as a delay in starting the task coupled with degraded memory, also matched the behaviour of children with ADHD. While these results are interesting, the neural/cognitive mechanisms reflected in the state regulation and cognitive deficit variables are unclear; and hence, the neurological roots of ADHD are no more apparent. However, it is

currently difficult to see how the results presented here, focussing on asymmetrically effective prediction errors, could give rise to the behaviour presented by Sonuga-Barke (2002a). Currently, my investigation of ADHD has focused on reinforcement-learning deficits in ADHD; however, the matching task is not a learning task *per se*. Hence, a more detail investigation of the mechanisms involved in this task could prove to be particularly enlightening as it may reveal additional dysfunctional systems associated with ADHD and how they interact with asymmetrical prediction errors.

5.1.2 Concluding remarks

In this thesis I have outlined a model-based analysis of the dopamine's role in ADHD. The central result that emerged from this analysis offers evidence that ADHD is the result of asymmetrical prediction error efficiency such that positive prediction error signals are more effective than negative signals. Though an exhaustive search of all the possible mechanisms through which this asymmetry may arise is beyond the scope of this work, I have outlined several of the most likely biological mechanisms involved in ADHD. Furthermore, an analysis of two existing model's of ADHD suggests that both could potentially improve their predictions of ADHD behaviour by incorporating these biological parameters. While this work in no way specifies the precise neurological mechanisms give rise to ADHD, it narrows the scope of potential dysfunctions considerably while also suggesting several fruitful paths forward in pursuit of a better understanding of both the disorder and the neurological systems involved.

Bibliography

Alsop, B. (2007). Problems with spontaneously hypertensive rats (SHR) as a model of attention-deficit/hyperactivity disorder (AD/HD). *Journal of Neuroscience Methods, 162*(1-2), 42-48.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (DSM-IV-TR)* American Psychiatric Association Washington, DC.

Arnsten, A. F., Steere, J. C., & Hunt, R. D. (1996). The contribution of alpha 2-noradrenergic mechanisms of prefrontal cortical cognitive function. potential significance for attention-deficit hyperactivity disorder. *Archives of General Psychiatry, 53*(5), 448-455.

Barkley, R. A. (1997a). *ADHD and the nature of self-control* The Guilford Press.

Barkley, R. A. (1997b). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*, 65-94.

Bellgrove, M. A., Domschke, K., Hawi, Z., Kirley, A., Mullins, C., Robertson, I. H., & Gill, M. (2005). The methionine allele of the COMT polymorphism impairs prefrontal cognition in children and adolescents with ADHD. *Experimental Brain Research, 163*(3), 352-360.

Berger, D. F., & Sagvolden, T. (1998). Sex differences in operant discrimination behaviour in an animal model of attention-deficit hyperactivity disorder. *Behavioural Brain Research, 94*(1), 73-82.

Bertsekas, D. P., & Tsitsiklis, J. N. (1996). *Neuro-dynamic programming* Athena Scientific.

Boix, F., Qiao, S., Kolpus, T., & Sagvolden, T. (1998). Chronic L-deprenyl treatment alters brain monoamine levels and reduces impulsiveness in an animal model of attention-deficit/hyperactivity disorder. *Behavioural Brain Research, 94*(1), 153-162.

Bradley, C. (1937). The behavior of children receiving benzedrine. *American Journal of Psychiatry, 94*(3), 577-585.

Bush, G., Frazier, J. A., Rauch, S. L., Seidman, L. J., Whalen, P. J., Jenike, M. A., Rosen, B. R., & Biederman, J. (1999). Anterior cingulate cortex dysfunction in attention-deficit/hyperactivity disorder revealed by fMRI and the counting stroop. *Biological Psychiatry, 45*(12), 1542-1552.

Casey, B. J., Castellanos, F. X., Giedd, J. N., Marsh, W. L., Hamburger, S. D., Schubert, A. B., Vauss, Y. C., Vaituzis, A. C., Dickstein, D. P., & Sarfatti, S. E. (1997). Implication of right frontostriatal circuitry in response inhibition and attention-deficit/hyperactivity disorder. *Journal of Amer Academy of Child & Adolescent Psychiatry, 36*(3), 374.

Castellanos, F. X. (1997). Toward a pathophysiology of attention-deficit/hyperactivint disorder. *Clinical Pediatrics, 36*(7), 381.

Castellanos, F. X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nature Reviews.Neuroscience, 3*(8), 617.

Catania, A. C., Sagvolden, T., & Keller, K. J. (1988). Reinforcement schedules: Retroactive and proactive effects of reinforcers inserted into fixed-interval performances. *Journal of the Experimental Analysis of Behavior, 49*(1), 49.

Chelune, G. J., Ferguson, W., Koon, R., & Dickey, T. O. (1986). Frontal lobe disinhibition in attention deficit disorder. *Child Psychiatry and Human Development, 16*(4), 221-234.

Coghill, D., Nigg, J., Rothenberger, A., Sonuga-Barke, E., & Tannock, R. (2005). Whither causal models in the neuroscience of ADHD? *Developmental Science, 8*(2), 105-114.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99*(1), 45-77.

Cooper, J. R., Bloom, F. E., & Roth, R. H. (2003). *The biochemical basis of neuropharmacology* Oxford University Press, USA.

Cragg, S. J., & Rice, M. E. (2004). DAncing past the DAT at a DA synapse. *Trends in Neurosciences, 27*(5), 270-277.

Dayan, P., & Niv, Y. (2008). Reinforcement learning: The good, the bad and the ugly. *Current Opinion in Neurobiology, 18*(2), 185-196.

Dayan, P., & Sejnowski, T. J. (1994). TD (X) converges with probability 1. *Machine Learning, 14*, 295-301.

Egelman, D. M., Person, C., & Montague, P. R. (1998). A computational role for dopamine delivery in human decision-making. *Journal of Cognitive Neuroscience, 10*(5), 623-630.

Floresco, S. B., West, A. R., Ash, B., Moore, H., & Grace, A. A. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nature Neuroscience, 6*, 968-973.

Frank, M. J., & O'Reilly, R. C. (2006). A mechanistic account of striatal dopamine function in human cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behavioral Neuroscience, 120*(3), 497-517.

Frank, M. J., Santamaria, A., O'Reilly, R. C., & Willcutt, E. (2006). Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology, 32*(7), 1583-1599.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron, 47*(4), 495-501.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated parkinsonism. *Journal of Cognitive Neuroscience, 17*(1), 51-72.

Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychological Review, 113*(2), 300-326.

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science, 306*(5703), 1940-1943.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107*(2), 289-344.

Goldman-Rakic, P. S., Leranth, C., Williams, S. M., Mons, N., & Geffard, M. (1989). Dopamine synaptic complex with pyramidal neurons in primate cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America, 86*(22), 9015-9019.

Gorenstein, E. E., Mammato, C. A., & Sandy, J. M. (1989). Performance of inattentive-overactive children on selected measures of prefrontal-type function. *Journal of Clinical Psychology, 45*(4)

Grace, A. A. (1991). Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: A hypothesis for the etiology of schizophrenia. *Neuroscience, 41*(1), 1.

Grace, A. A. (2001). Psychostimulant actions on dopamine and limbic system function: Relevance to the pathophysiology and treatment of ADHD. *Stimulant Drugs and ADHD: Basic and Clinical Neuroscience,* , 134–157.

Gray, J. A., Feldon, J., Rawlins, J. N. P., Hemsley, D. R., & Smith, A. D. (1991). The neuropsychology of schizophrenia. *Behavioral and Brain Sciences, 14*(1), 1-84.

Grodzinsky, G. M., Diamond, R. (1992). Frontal lobe functioning in boys with attention-deficit hyperactivity disorder. *Developmental Neuropsychology, 8*(4), 427-445.

Hernandez-Lopez, S., Bargas, J., Surmeier, D. J., Reyes, A., & Galarraga, E. (1997). D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca2 conductance. *Journal of Neuroscience, 17*(9), 3334.

Herrnstein, R. J. (1991). Experiments on stable suboptimality in individual behavior. *The American Economic Review,* , 360-364.

Holroyd, C. B., Baker, T. E., Kerns, K. A., & Müller, U. (2008). Electrophysiological evidence of atypical motivation and reward processing in children with attention-deficit hyperactivity disorder. *Neuropsychologia,*

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679-709.

Krause, K. H., Dresel, S. H., Krause, J., la Fougere, C., & Ackenheil, M. (2003). The dopamine transporter and neuroimaging in attention deficit hyperactivity disorder. *Neuroscience and Biobehavioral Reviews, 27*(7), 605-613.

Lou, H. C., Henriksen, L., Bruhn, P., Borner, H., & Nielsen, J. B. (1989). Striatal dysfunction in attention deficit and hyperkinetic disorder. *Archives of Neurology, 46*(1), 48-52.

Luman, M., Oosterlaan, J., & Sergeant, J. A. (2005). The impact of reinforcement contingencies on AD/HD: A review and theoretical appraisal. *Clinical Psychology Review, 25*(2), 183-213.

Madras, B. K., Miller, G. M., & Fischman, A. J. (2002). The dopamine transporter: Relevance to attention deficit hyperactivity disorder (ADHD). *Behavioural Brain Research, 130*(1-2), 57-63.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*(5695), 503.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the rescorla-wagner model. *Psychological Bulletin, 117*(3), 363-386.

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology, 50*(4), 381-425.

Missale, C., Nash, S. R., Robinson, S. W., Jaber, M., & Caron, M. G. (1998). Dopamine receptors: From structure to function. *Physiological Reviews, 78*(1), 189-225.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of Neuroscience, 16*(5), 1936-1947.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology,*

Nowycky, M. C., & Roth, R. H. (1977). Presynaptic dopamine receptors. *Naunyn-Schmiedeberg's Archives of Pharmacology, 300*(3), 247-254.

O'Donnell, P., & Grace, A. A. (1995). Synaptic interactions among excitatory afferents to nucleus accumbens neurons: Hippocampal gating of prefrontal cortical input. *Journal of Neuroscience, 15*(5), 3622.

Pedarzani, P., & Storm, J. F. (1995). Dopamine modulates the slow ca (2)-activated K current IAHP via cyclic AMP-dependent protein kinase in hippocampal neurons. *Journal of Neurophysiology, 74*(6), 2749-2753.

Pucak, M. L., & Grace, A. A. (1994). Evidence that systemically administered dopamine antagonists activate dopamine neuron firing primarily by blockade of somatodendritic autoreceptors. *Journal of Pharmacology and Experimental Therapeutics, 271*(3), 1181-1192.

Redish, A. D., Jensen, S., Johnson, A., & Kurth-Nelson, Z. (2007). Reconciling reinforcement learning models with behavioral extinction and renewal: Implications for addiction, relapse, and problem gambling. *Psychological Review, 114*(3), 784-805.

Redish, A. D. (2004). *Addiction as a computational process gone awry* American Association for the Advancement of Science.

Richfield, E. K., Penney, J. B., & Young, A. B. (1989). Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. *Neuroscience, 30*(3), 767-777.

Rubia, K., Overmeyer, S., Taylor, E., Brammer, M., Williams, S. C. R., Simmons, A., & Bullmore, E. T. (1999). Hypofrontality in attention deficit hyperactivity disorder during higher-order motor control: A study with functional MRI. *American Journal of Psychiatry, 156*(6), 891-896.

Russell, V. A., Oades, R. D., Tannock, R., Killeen, P. R., Auerbach, J. G., Johansen, E. B., & Sagvolden, T. (2006). Response variability in attention-Deficit/Hyperactivity disorder: A neuronal and glial energetics hypothesis. *Behavioral and Brain Functions : BBF, 2*, 30. doi:10.1186/1744-9081-2-30

Sagvolden, T. (2000). Behavioral validation of the spontaneously hypertensive rat (SHR) as an animal model of attention-deficit/hyperactivity disorder (AD/HD). *Neuroscience and Biobehavioral Reviews, 24*(1), 31-39.

Sagvolden, T., Aase, H., Zeiner, P., & Berger, D. (1998). Altered reinforcement mechanisms in attention-deficit/hyperactivity disorder. *Behavioural Brain Research, 94*(1), 61-71.

Sagvolden, T., Hendley, E. D., & Knardahl, S. (1992). Behavior of hypertensive and hyperactive rat strains: Hyperactivity is not unitarily determined. *Physiology & Behavior, 52*(1), 49-57.

Sagvolden, T., Johansen, E. B., Aase, H., & Russell, V. A. (2005). A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes. *Behavioral and Brain Sciences, 28*(03), 397-419.

Sagvolden, T., Pettersen, M. B., & Larsen, M. C. (1993). Spontaneously hypertensive rats (SHR) as a putative animal model of childhood hyperkinesis: SHR behavior compared to four other rat strains. *Physiology & Behavior, 54*(6), 1047-1055.

Sagvolden, T., Russell, V. A., Aase, H., Johansen, E. B., & Farshbaf, M. (2005). Rodent models of attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57*(11), 1239-1247.

Sajikumar, S., & Frey, J. U. (2004). Late-associativity, synaptic tagging, and the role of dopamine during LTP and LTD. *Neurobiology of Learning and Memory, 82*(1), 12-25.

Scahill, L., & Schwab-Stone, M. (2000). Epidemiology of ADHD in school-age children. *Child and Adolescent Psychiatric Clinics of North America, 9*(3), 541-55, vii.

Schachar, R. J., Chen, S., Logan, G. D., Ornstein, T. J., Crosbie, J., Ickowicz, A., & Pakulak, A. (2004). Evidence for an error monitoring deficit in attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology, 32*(3), 285-293.

Scheres, A., Lee, A., & Sumiya, M. (2008). Temporal reward discounting and ADHD: Task and symptom specific effects. *Journal of Neural Transmission, 115*(2), 221-226.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36*(2), 241-263.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80*(1), 1-27.

Sergeant, J. A. (2005). Modeling attention-deficit/hyperactivity disorder: A critical appraisal of the cognitive-energetic model. *Biological Psychiatry, 57*(11), 1248-1255.

Sergeant, J. A., & van der Meere, J. (1988). What happens after a hyperactive child commits an error? *Psychiatry Research, 24*(2), 157-164.

Servan-Schreiber, D., Printz, H., & Cohen, J. (1990). A network model of catecholamine effects: Gain, signal-to-noise ratio, and behavior. *Science, 249*(4971), 892-895.

Sieg, K. G., Gaffney, G. R., Preston, D. F., & Hellings, J. A. (1995). SPECT brain imaging abnormalities in attention deficit hyperactivity disorder. *Clinical Nuclear Medicine, 20*(1), 55.

Sleator, E. K., & Ullmann, R. K. (1981). Can the physician diagnose hyperactivity in the office? *Pediatrics, 67*(1), 13-17.

Smith, A. D., & Bolam, J. P. (1990). The neural network of the basal ganglia as revealed by the study of synaptic connections of identified neurones. *Trends in Neurosciences, 13*(7), 259-265.

Solanto, M. V. (1990). The effects of reinforcement and response-cost on a delayed response task in children with attention deficit hyperactivity disorder: A research note. *Journal of Child Psychology and Psychiatry, 31*(5), 803-808.

Sonuga-Barke, E. J. S. (2002a). Interval length and time-use by children with AD/HD: A comparison of four models. *Journal of Abnormal Child Psychology, 30*(3), 257-264.

Sonuga-Barke, E. J. S. (2002b). Psychological heterogeneity in AD/HD—a dual pathway model of behaviour and cognition. *Behavioural Brain Research, 130*(1-2), 29-36.

Sonuga-Barke, E. J. S., Taylor, E., Sembi, S., & Smith, J. (1992). Hyperactivity and delay aversion-I. the effect of delay on choice. *Journal of Child Psychology and Psychiatry, 33*(2), 387-398.

Sutker, P. B., & Adams, H. E. (2001). *Comprehensive handbook of psychopathology(3rd ed)* Plenum press New York.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning, 3*(1), 9-44.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review, 88*(2), 135-170.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* MIT Press Cambridge, MA.

Swanson, J. M., Flodman, P., Kennedy, J., Spence, M. A., Moyzis, R., Schuck, S., Murias, M., Moriarity, J., Barr, C., & Smith, M. (2000). Dopamine genes and ADHD. *Neuroscience and Biobehavioral Reviews, 24*(1), 21-25.

Taylor, E. (1994). Syndromes of attention deficit and overactivity. *Child and Adolescent Psychiatry.Modern Approaches,* , 285–307.

Tripp, G., & Alsop, B. (1999). Sensitivity to reward frequency in boys with attention deficit hyperactivity disorder. *Journal of Clinical Child & Adolescent Psychology, 28*(3), 366-375.

Tripp, G., & Wickens, J. (2008). Research review: Dopamine transfer deficit: A neurobiological theory of altered reinforcement mechanisms in ADHD. *Journal of Child Psychology and Psychiatry, 49*(7), 691-704.

Vitiello, B., Severe, J. B., Greenhill, L. L., Arnold, L. E., Abikoff, H. B., Bukstein, O. G., Elliott, G. R., Hechtman, L., Jensen, P. S., Hinshaw, S. P., March, J. S., Newcorn, J. H., Swanson, J. M., & Cantwell, D. P. (2001). Methylphenidate dosage for children with ADHD over time under controlled conditions: Lessons from the MTA. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(2), 188-196.

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature, 412*(6842), 43-48.

Wilens, T. E., Biederman, J., & Spencer, T. J. (2002). Attention deficit/hyperactivity disorder across the lifespan. *Annual Review of Medicine, 53*(1), 113-131.

Wilens, T. E., Faraone, S. V., & Biederman, J. (2004). Attention-deficit/hyperactivity disorder in adults. *JAMA, 292*(5), 619-623.

Williams, J., & Dayan, P. (2005). Dopamine, learning, and impulsivity: A biological account of attention-Deficit/Hyperactivity disorder. *Journal of Child & Adolescent Psychopharmacology, 15*(2), 160-179.

Williams, J., & Taylor, E. (2004). Dopamine appetite and cognitive impairment in attention Deficit/Hyperactivity disorder. *Neural Plasticity, 11*(1)

Wilson, C. J., & Kawaguchi, Y. (1996). The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. *Journal of Neuroscience, 16*(7), 2397.

Wultz, B., & Sagvolden, T. (1992). The hyperactive spontaneously hypertensive rat learns to sit still, but not to stop bursts of responses with short interresponse times. *Behavior Genetics, 22*(4), 415-433.

Zametkin, A. J., Liebenauer, L. L., Fitzgerald, G. A., King, A. C., Minkunas, D. V., Herscovitch, P., Yamada, E. M., & Cohen, R. M. (1993). Brain metabolism in

teenagers with attention-deficit hyperactivity disorder. *Archives of General Psychiatry, 50*(5), 333.

Zametkin, A. J., Nordahl, T. E., Gross, M., & King, A. C. (1990). Cerebral glucose metabolism in adults with hyperactivity of childhood onset. *New England Journal of Medicine, 323*(20), 1361-1366.