Testing Computational Models of Dopamine and Noradrenaline Dysfunction in Attention Deficit/Hyperactivity Disorder

Michael J Frank*[1], Amy Santamaria[2,3], Randall C O'Reilly[2] and Erik Willcutt[2]

[1]Department of Psychology and Program in Neuroscience, University of Arizona, Tucson, AZ, USA; [2]Department of Psychology, University of Colorado at Boulder, Boulder, CO, USA; [3]Alion Science and Technology, MA&D Operation, Boulder, CO, USA

We test our neurocomputational model of fronto-striatal dopamine (DA) and noradrenaline (NA) function for understanding cognitive and motivational deficits in attention deficit/hyperactivity disorder (ADHD). Our model predicts that low striatal DA levels in ADHD should lead to deficits in 'Go' learning from positive reinforcement, which should be alleviated by stimulant medications, as observed with DA manipulations in other populations. Indeed, while nonmedicated adult ADHD participants were impaired at both positive (Go) and negative (NoGo) reinforcement learning, only the former deficits were ameliorated by medication. We also found evidence for our model's extension of the same striatal DA mechanisms to working memory, via interactions with prefrontal cortex. In a modified AX-continuous performance task, ADHD participants showed reduced sensitivity to working memory contextual information, despite no global performance deficits, and were more susceptible to the influence of distractor stimuli presented during the delay. These effects were reversed with stimulant medications. Moreover, the tendency for medications to improve Go relative to NoGo reinforcement learning was predictive of their improvement in working memory in distracting conditions, suggestive of common DA mechanisms and supporting a unified account of DA function in ADHD. However, other ADHD effects such as erratic trial-to-trial switching and reaction time variability are not accounted for by model DA mechanisms, and are instead consistent with cortical noradrenergic dysfunction and associated computational models. Accordingly, putative NA deficits were correlated with each other and independent of putative DA-related deficits. Taken together, our results demonstrate the usefulness of computational approaches for understanding cognitive deficits in ADHD.

INTRODUCTION

Attention deficit/hyperactivity disorder (ADHD) is a common childhood-onset psychiatric condition characterized by age-inappropriate levels of inattention and/or hyperactivity-impulsivity (APA, 1994). Despite widespread public skepticism regarding the legitimacy of ADHD as a disorder, several recent findings demonstrate clear biological underpinnings. These findings include multiple genetic factors, ADHD-related differences in brain structure and function, and changes in neurotransmitter components within the fronto-striatal system (for recent reviews, see Castellanos *et al*, 2006; Krain and Castellanos, 2006; Faraone *et al*, 2005). The main functional deficits of ADHD are less clear. Neuropsychological studies reveal executive function deficits, with particularly reliable impairments in response inhibition (Willcutt *et al*, 2005). However, the modest effect sizes in these studies suggest that executive dysfunction is neither necessary nor sufficient to account for ADHD symptoms. Other findings point to key deficits in motivational/reward processes (Luman *et al*, 2005; Sagvolden *et al*, 1998; Aase *et al*, 2006) and suggest that these are largely independent of response inhibition deficits (Solanto *et al*, 2001; Toplak *et al*, 2005). These findings have led to a 'dual-pathway' hypothesis for ADHD. In one pathway, executive function and response inhibition deficits result from impaired circuitry linking dorsal striatum and dorsolateral prefrontal cortex. In the second pathway, motivational deficits result from reduced processing in ventral striatal-orbitofrontal circuits (Sonuga-Barke, 2003; Castellanos *et al*, 2006).

*Correspondence: Dr MJ Frank, Department of Psychology and Program in Neuroscience, University of Arizona, 1503 E University Blvd, Bldg #68, Tucson, AZ 85721, USA, Tel: +1 520 626 4787; Fax: +1 520 621 9306, E-mail: mfrank@u.arizona.edu

Figure 1 (a) The striato-cortical loops, including the direct ('Go') and indirect ('NoGo') pathways of the BG. Striatal Go cells disinhibit the thalamus via GPi, thereby facilitating the execution of an action represented in cortex. The NoGo cells have an opposing effect by increasing inhibition of the thalamus, suppressing actions from getting executed. Dopamine from the SNc projects to the dorsal striatum, causing excitation of Go cells via D1 receptors, and inhibition of NoGo via D2 receptors. Noradrenaline from the LC modulates activity in frontal cortex. GPi: internal segment of globus pallidus; GPe: external segment of globus pallidus; SNc: substantia nigra pars compacta; STN: subthalamic nucleus; LC: locus coeruleus. (b) Our neural network model of this circuit (Frank, 2005; Frank *et al*, in press), where squares represent units, with height reflecting neural activity. The Premotor Cortex selects an Output response via direct projections from the sensory Input, and is modulated by the BG projections from Thalamus. Go units are in the left half of the Striatum layer; NoGo in the right half, with separate columns for the two responses (R1 and R2). In the case shown, striatum Go is stronger than NoGo for R1, inhibiting GPi, disinhibiting Thalamus, and facilitating R1 execution in cortex. A tonic level of dopamine is shown in SNc; a burst or dip ensues in a subsequent reinforcement phase (not shown), driving Go/NoGo learning. The STN exerts a dynamic 'Global NoGo' function on the execution of all responses, complementing the response-specific striatal NoGo cells (Frank, 2006). The LC dynamically modulates the gain of premotor units, increasing the signal-to-noise ratio and modulating variability in reaction times. Similar models have been used to simulate interactions between the BG and prefrontal cortices in working memory and decision making (not shown; O'Reilly and Frank, 2006; Frank and Claus, 2006).

In this paper, we take a somewhat different approach, motivated by principles developed within neurocomputational models of cognition. We use this computational framework to make predictions about the mechanisms that produce cognitive and motivational deficits in ADHD. In particular, our models simulate dynamic dopamine (DA) interactions within circuits linking the basal ganglia (BG) with frontal cortex (Alexander *et al*, 1986) and explore their implication in action selection, reinforcement learning, working memory, and decision making (Frank *et al*, 2001; Frank, 2005, 2006; O'Reilly and Frank, 2006; Frank and Claus, 2006) (for details about the models, see Figure 1 and the articles cited above). These models are explicit about the nature of neural interactions and demonstrate how cognitive and motivational processes can emerge from network dynamics. They have provided insight into the source of cognitive deficits in patients with Parkinson's disease and those with orbitofrontal damage (Frank, 2005; Frank and Claus, 2006), making novel predictions that have been confirmed in experiments with medicated and nonmedicated Parkinson's patients, experiments with healthy participants taking dopamine D2 agonists and antagonists, and electrophysiological studies (Frank *et al*, 2004, 2005; Frank and O'Reilly, 2006). Here we test model predictions in ADHD participants on and off stimulant medications in reinforcement learning and working memory tasks. We show that a single mechanism (reduced DA in the striatum) can account for both motivational and working memory deficits, while another (likely noradrenergic) mechanism may account for other behaviorally independent aspects of the disorder.

Dopamine Dysfunction in ADHD

There is a growing body of evidence that ADHD is associated with low levels of striatal DA (eg, Sagvolden *et al*, 2005; Biederman and Faraone, 2002; Solanto, 2002). In

a comprehensive review of the biological basis of ADHD, the authors concluded that decreased dopaminergic function in three separate striato-cortical loops led to deficits in reinforcement and extinction behaviors (Sagvolden *et al*, 2005). Reduced striatal DA has also been demonstrated in human ADHD participants. Both children and adults with ADHD have abnormally high densities of striatal dopamine transporters (DATs), so that too much DA is removed from the synapse (Dougherty *et al*, 1999; Krause *et al*, 2000). This finding is also supported by altered DAT genetic factors in ADHD (Faraone *et al*, 2005; Todd *et al*, 2005). There has been some disagreement about whether both tonic and phasic DA levels are reduced in ADHD. In fact, some early accounts argued that low tonic DA levels actually lead to *enhanced* phasic DA, due to reduced tonic DA stimulation onto inhibitory autoreceptors (Grace, 2001; Solanto, 2002). It was further argued that stimulant medications normalize DA dysfunction by enhancing tonic DA and decreasing phasic DA (Seeman and Madras, 2002). However, other data show that stimulants do not have preferential action on autoreceptors at any dose (Ruskin *et al*, 2001). The enhanced phasic DA hypothesis has since been rejected by some authors, who now believe that ADHD is associated with reduced levels of both tonic and phasic DA (Madras *et al*, 2005; Sagvolden *et al*, 2005). This view is supported by evidence that stimulants increase both extracellular striatal DA (Volkow *et al*, 2001), and synaptic DA associated with phasic responses (Schiffer *et al*, 2006). This medication-induced increase in phasic DA is thought to provide a focusing and teaching signal in ADHD (eg, Schultz, 2002). Furthermore, stimulants increase DA release to a greater extent in striatum than in prefrontal cortex (Mazei *et al*, 2002; Madras *et al*, 2005), likely due to the far greater DAT density in striatum (Sesack *et al*, 1998; Cragg *et al*, 2002). This suggests that phasic DA effects of ADHD and associated medications may be particularly relevant in the striatum. Nevertheless, increases in striatal activity can lead

to indirect effects in frontal cortex, via systems-level interactions (Alexander *et al*, 1986). Supporting this notion, stimulant induced enhancements in activity and plasticity in the striatal 'direct pathway' were highly correlated with increases in frontal cortical activity (Yano and Steiner, 2005).

The above evidence suggests that ADHD is associated with reduced tonic and phasic striatal DA levels, which are alleviated by stimulant medications. However, exactly how would these effects propagate through the system to affect behavior? Our computational models suggest that healthy levels of DA are required to dynamically modulate the balance of 'Go' and 'NoGo' pathways in the striatum. Via systems-level interactions, these pathways act to facilitate or suppress the execution of actions represented in frontal cortex. Actions range from lower-level motor programs (in premotor areas) to higher-level working memory updating and decision making (in dorsolateral prefrontal and orbitofrontal areas). In particular, phasic DA signals that occur during positive and negative reinforcement (eg, Schultz, 2002) drive learning in the models to facilitate the selection of rewarding actions and to prevent selection of those that are less rewarding (Frank, 2005). In the working memory domain, these DA signals drive updating of task-relevant information into dorsolateral prefrontal working memory representations (O'Reilly and Frank, 2006), and similarly facilitate the maintenance of long-term reward values in orbitofrontal areas for use in decision making (Frank and Claus, 2006). If striatal DA is reduced in ADHD, the models predict ADHD-related deficits in (i) positive (Go) reinforcement learning; (ii) selective updating of task-relevant information into working memory; and (iii) modulation of long-term reward information needed for decision making. Predictions for aversive (NoGo) reinforcement learning are somewhat more ambiguous, given that multiple brain mechanisms contribute to this learning, some of which may also be compromised in ADHD. Nevertheless, the models clearly predict that all the above deficits should be ameliorated by stimulant medications which increase DA signals and selectively enhance plasticity in striatal Go cells (eg, Yano and Steiner, 2005; Yano *et al*, 2006). Thus, we predicted that medications would selectively enhance Go but not NoGo learning in ADHD.

Noradrenaline Dysfunction in ADHD

Nevertheless, it is clear that ADHD is not a unitary disorder (Castellanos *et al*, 2006; Nigg and Casey, 2005; Diamond, 2005), and other behavioral symptoms of ADHD are not easily explained by the reduced DA hypothesis. For example, ADHD participants typically show more within-subject variability in their reaction times (Leth-Steensen *et al*, 2000; Lijffijt *et al*, 2005; Castellanos *et al*, 2005). Although DA modulates the overall reaction time in our models (by modulating Go *vs* NoGo pathways in the striatum; Frank, 2005), it is not immediately evident how low levels of DA would lead to more variability in RTs across trials (but see General Discussion for alternative accounts). In contrast, both computational and experimental work have shown that cortical noradrenaline (NA) signals do affect response variability (Usher *et al*, 1999; Aston-Jones and Cohen, 2005; Frank *et al*, in press). The

models show that phasic NA release leads to 'sharper' motor cortical representations and a tighter distribution of reaction times, whereas a high tonic but low phasic state is associated with more RT variability. Further, increases in tonic NA levels are thought to enable the representation of competing cortical representations during exploration of new behaviors, which may also lead to erratic behavior. In accord with these computational accounts, we have previously proposed that some aspects of response variability and response inhibition effects in ADHD can be accounted for by cortical NA dysfunction (Frank *et al*, in press; see also Zametkin and Rapoport, 1987; Arnsten *et al*, 1996; Aron and Poldrack, 2005). The NA hypothesis is supported by evidence for elevated frontal cortical NA levels in a rat model of ADHD (Russell and Wiggins, 2000; Russell *et al*, 2000). In human ADHD participants, RT variability is correlated with noradrenergic (and not dopaminergic) measures (Llorente *et al*, 2006). In addition, deficits in response inhibition are ameliorated by selective NA transporter blockers (eg, Swanson *et al*, 2006; Michelson *et al*, 2001; Overtoom *et al*, 2003; Chamberlain *et al*, 2006).

Summary and Predictions

In sum, both DA and NA effects may be critical for heterogeneous deficits in ADHD. It is plausible that DA effects are involved in motivation/reward and working memory updating, while NA effects are involved in response inhibition and variability, supporting the independence of these symptoms (Solanto *et al*, 2001). Based on the above computational considerations, we make the following predictions.

1. Deficits related to DA dysfunction (positive reinforcement learning and working memory updating) should be correlated with each other and should be similarly improved by medications that increase DA.
2. Deficits related to NA dysfunction (reaction time variability and erratic trial-to-trial exploratory behavior) should be correlated with each other but should not be improved by medications that increase DA.
3. These two kinds of deficits should be independent of one another.

We tested these predictions in adults with ADHD and healthy control participants. ADHD participants were tested both on and off standard stimulant medications which increase both DA and NA levels (Madras *et al*, 2005). In our experiments, we tested two tasks: probabilistic selection (two alternative forced-choice), and modified versions of the widely-used AX-CPT working memory task. The probabilistic selection task has been used previously to test model predictions about individual differences in learning from positive *vs* negative reinforcement in Parkinson patients (Frank *et al*, 2004), healthy participants taking DA agonists and antagonists (Frank and O'Reilly, 2006) and electrophysiological correlates (Frank *et al*, 2005). The AX-CPT working memory task (Servan-Schreiber *et al*, 1997; Barch *et al*, 1997, 2001) is used to test whether the same Go/NoGo processes at work in the simpler procedural learning tasks also apply to working memory updating, as predicted by the models and supported by pharmacological

manipulations in healthy individuals (Frank and O'Reilly, 2006). Importantly, our paradigms enable analysis that goes beyond global impairment, by examining *relative* within-subject differences between conditions that should be differentially affected by DA. This approach is critical for making specific predictions that are not susceptible to the explanation of ADHD subjects simply not paying attention to the overall task rules, which could potentially explain deficits across a wide range of tasks.

EXPERIMENT

Reinforcement Learning Task

In the *Probabilistic Selection* task (Frank *et al*, 2004, 2005; Frank and O'Reilly, 2006), three different stimulus pairs (AB, CD, EF) are presented in random order and participants have to learn to choose one of the two stimuli (Figure 2). Feedback follows the choice to indicate whether it was correct or incorrect, but this feedback is probabilistic. In AB trials, a choice of stimulus A leads to correct (positive) feedback in 80% of AB trials, whereas a B choice leads to incorrect (negative) feedback in these trials (and vice versa for the remaining 20% of trials). CD and EF pairs are less reliable: stimulus C is correct in 70% of CD trials, while E is correct in 60% of EF trials. Over the course of training, participants learn to choose stimuli A, C, and E more often than B, D, or F. Note that learning to choose A over B could be accomplished either by learning that A leads to positive feedback or that B leads to negative feedback (or both). To evaluate whether participants learned more about positive or negative outcomes of their decisions, we subsequently tested them with novel combinations of stimulus pairs involving either an A (AC, AD, AE, AF) or a B (BC, BD, BE, BF); no feedback was provided. If

participants had learned more from positive feedback, they should reliably choose stimulus A in all novel test pairs in which it is present. On the other hand, if they learned more from negative feedback, they should more reliably avoid stimulus B. Indeed, we recently showed that avoidance of stimulus B is associated with enhanced negative feedback related brain potentials (Frank *et al*, 2005).

We hypothesized that, due to low striatal phasic DA signals, ADHD participants would be impaired at positive Go reinforcement learning and that this deficit would be normalized by stimulant medications that increase DA. These predictions are based on our models, DA manipulations in other populations, and recent neuroimaging data showing reduced striatal activation in ADHD during anticipation of rewards (Scheres *et al*, 2006b). It is important to note that we predict specific improvements in Go learning. We predicted no improvements in NoGo learning because, according to our model, low DA levels are needed to learn NoGo responses (Frank, 2005). In contrast, if improvements with medication are simply due to increasing vigilance or attention, medication should improve performance on both Go and NoGo learning. A specific improvement on Go learning would have both theoretical and practical implications because it would suggest that ADHD participants on medication respond more to positive incentives, and that negative reinforcement or punishments may be less effective.

We also collected reaction time data in light of previous findings that ADHD participants show more within-subjects variability in reaction times (Leth-Steensen *et al*, 2000; Castellanos *et al*, 2005). We hypothesized that the mechanism for this variability is high tonic NA and that this would also lead to more trial-to-trial erratic choice behavior during training. We further predicted that this variability would not be correlated with putative DA-related mechanisms.
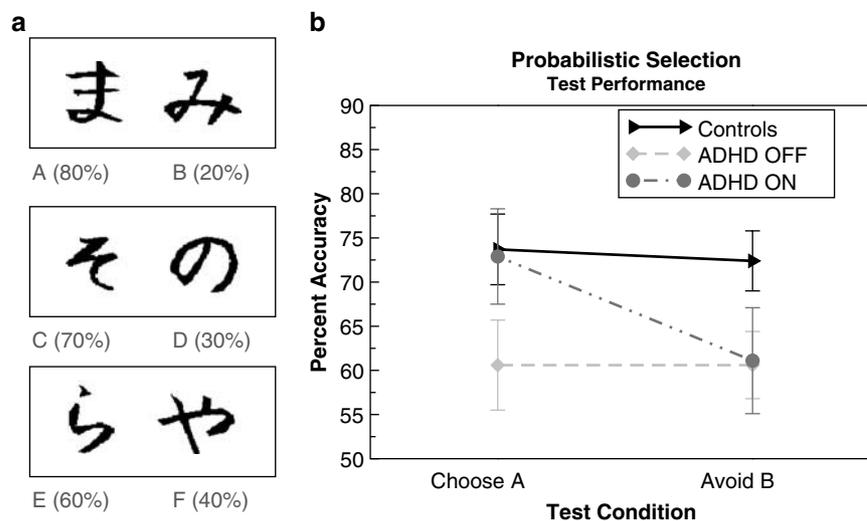
Figure 2 Probabilistic selection task and test results. (a) Example stimulus pairs (Japanese Hiragana characters), which minimize explicit verbal encoding. Each pair is presented separately in different trials. Three different pairs are presented in random order; correct choices are determined probabilistically, with the probability of positive reinforcement for selecting each stimulus shown in parentheses; negative (incorrect) reinforcement was delivered otherwise. (b) Novel test pair results, where choosing A depends on having learned from positive reinforcement and avoiding B depends on having learned to avoid negative reinforcement. As in previous studies, healthy performed equivalently in the two conditions. ADHD was associated with nonselective impairments in the test phase. Medication selectively improved positive reinforcement learning, while having no effect on avoid-B performance, consistent with increases in phasic DA in the basal ganglia. Error bars reflect standard error of the mean.

Results and discussion: putative DA effects. Results were broadly consistent with these predictions (Figure 2). Overall, there was no main effect of positive/negative test pair condition ($F(1, 35) = 2.2$, NS). Relative to controls, OFF ADHD participants were globally impaired at choosing among novel test pairs ($F(1, 35) = 6.2$, $p = 0.017$), with no test pair condition interaction ($F(1,35) = 0.2$). In particular, OFF ADHD participants were impaired at both positive (choose-A; $F(1, 35) = 4.9$, $p = 0.03$) and negative (avoid-B; $F(1, 25) = 3.7$, $p = 0.06$) pairs. In contrast, ON ADHD participants were unimpaired relative to controls ($F(1, 35) = 1.6$, NS), and there was no interaction with test pair condition ($F(1, 35) = 1.3$, NS). The main effect of stimulant medication on test performance was not significant ($F(1, 35) = 2.1$, NS). Nevertheless, further planned comparisons revealed that medication significantly improved positive/Go learning ($F(1, 35) = 4.1$, $p = 0.05$; effect size $d = 0.51$), such that choose-A performance in ON participants did not differ from controls ($F(1, 35) = 0.18$; effect size $d = 0.03$). There was no effect of medication on NoGo learning in ADHD ($F(1, 35) = 0.03$; $d = 0.02$). The interaction between ADHD deficits and positive/negative condition did not reach significance ($F(1, 35) = 2.1$, $p = 0.15$). [A significant interaction requires the relative within subject difference in Go *vs* NoGo learning to depend on medication status. While this difference was on average precisely zero in OFF participants (compared with 12% difference in ON participants), the inclusion of poor performers contributed substantial variability to this measure. Indeed, after excluding participants who did not perform better than chance levels during the test phase (this filter excluded six OFF ADHD participants, two ON, and five single sessions from controls), the interaction was significant ($F(1, 33) = 5.0$, $p = 0.03$).]

The improvement in positive Go learning by stimulant medication is consistent with predictions from our BG/DA model, and is similar to patterns we observed in Parkinson's patients on DA medication (Frank *et al*, 2004), and in healthy participants taking haloperidol, which at single low doses increases phasic DA release (Frank and O'Reilly, 2006; Wu *et al*, 2002). Our model further predicts that by elevating tonic DA levels, stimulant medications may block

the effects of DA dips and could therefore even impair NoGo learning (Frank, 2005), as seen with DA medications in several other populations (see Frank and O'Reilly, 2006 and references therein). This prediction was not fully borne out, in that ON participants were not worse than OFF participants at NoGo learning. Nevertheless, NoGo learning in ON participants was still relatively impaired compared with controls ($F(1.35) = 3.3$, $p = 0.08$; $d = 0.526$), and in fact was only marginally better than chance (50%) levels ($t(17) = 1.85$, $p = 0.08$), making it difficult for them to perform much worse. Thus, it remains possible that deleterious effects of stimulant medication on NoGo learning in ADHD were obscured by a floor effect. Alternatively it is possible that the medications do not sufficiently elevate tonic DA levels to effectively block NoGo learning. Consistent with this hypothesis, recent rat studies showed that stimulants selectively increased activity and gene expression in striatal Go cells, with no effect on NoGo cells (Yano and Steiner, 2005).

To further investigate the sources of impaired test performance in ADHD, we examined performance in the training phase. There was no main effect of group (ADHD *vs* controls) on the number of training trials needed to advance to the test phase ($F(1, 35) = 1.5$). Nevertheless, planned comparisons revealed that OFF ADHD participants required significantly more training trials ($M = 258$) than controls ($M = 177$; $F(1, 35) = 9.5$, $p = 0.004$). Medication sped up learning ($F(1, 35) = 9.9$, $p = 0.003$), such that medicated participants ($M = 173$) did not differ from controls ($F(1, 35) = 0.02$). Slowed learning in OFF ADHD participants was also reflected in reduced accuracy across training compared with controls ($F(1, 35) = 9.5$, $p = 0.004$; Figure 3a). Again, participants ON medication were unimpaired ($F(1, 35) = 0.2$), and the within-subject medication effect was significant ($F(1, 35) = 8.5$, $p = 0.006$). Moreover, both controls and medicated participants differentiated between reinforcement values of the training pairs (AB, CD, EF), whereas ADHD participants OFF medication did not (Figure 3b). Specifically, a within-subjects ANOVA in the last block of training revealed a main effect of training pair condition (AB, CD, EF) for controls ($F(2, 74) = 7.8$, $p < 0.001$) and ADHD participants ON
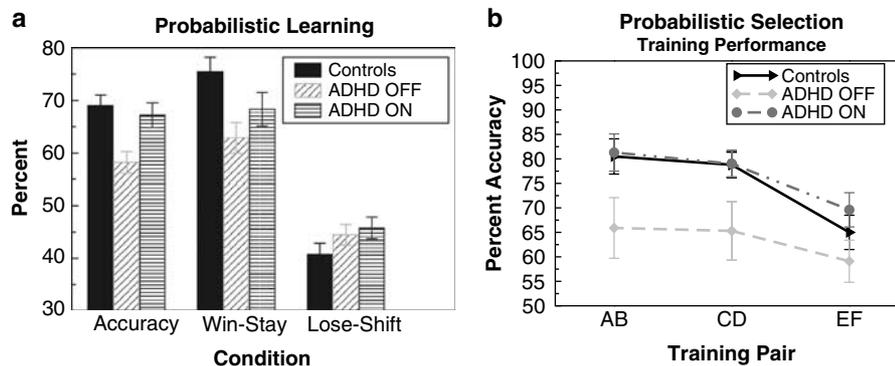

Figure 3 Probabilistic Selection training performance. (a) Overall accuracy during training was impaired in ADHD, and improved by medication. Impaired accuracy in OFF ADHD participants was also associated with reduced tendency to choose the same stimulus after having been rewarded for choosing it on the most recent trial of the same type (Win-Stay). No significant group differences were seen across in the tendency to modify choice behavior following negative feedback (Lose-Shift). The nonsignificant increase in ADHD lose-shift measure was due to their more erratic behavior overall, switching their choice behavior from trial to trial, independent of feedback (Figure 4). (b) Performance accuracy in the last block of training, broken down into the three training pairs. ADHD participants OFF medication showed an inability to discriminate between the different contingencies; medication normalized performance.

medication (F(2, 34) = 5.1, p = 0.01), but not those OFF medication (F(2, 32) = 0.5). Thus, the impaired generalization of reinforcement values in the test phase could be explained by deficits in learning of the original contingencies in ADHD, as would be expected by reduced DA bursts.

To examine Go *vs* NoGo learning during training, we analyzed trial-to-trial behavior as a function of positive and negative feedback. There was a main effect of group, such that following positive feedback, ADHD participants were significantly less likely than controls to choose the same stimulus in the next trial in which it appeared (win-stay; Figure 3a; F(1, 35) = 4.9, p = 0.03). Planned comparisons revealed that this effect was significant in OFF patients (F(1, 35) = 7.1, p = 0.01), whereas those ON medication were unimpaired in this measure (F(1, 35) = 2.3, NS). Conversely, following negative feedback, there was no effect of group on lose-shift performance (F(1, 35) = 1.8, NS). Participants OFF medication were just as likely as controls to switch choices in the subsequent trial of the same type (F(1, 35) = 1.0); there was no effect of medication on this measure (F(1, 35) = 0.2). Overall, these analyses are consistent with a selective impairment in positive feedback learning depending on phasic DA in ADHD.

Results and discussion: putative noradrenaline effects. Why were ADHD participants impaired at avoid-B during test, when they showed intact switching following negative feedback during training? It is possible that intact switching after negative feedback in ADHD participants was an artifact of an overall greater tendency to switch, regardless of the previous trial's feedback. Such erratic switching behavior is theoretically predicted by models of noradrenergic dysfunction (eg, Aston-Jones and Cohen, 2005; McClure *et al*, 2006; Frank *et al*, in press), as discussed above. Indeed, we found that ADHD participants switched their responses from trial to trial more often than controls (F(1, 35) = 4.24, p = 0.047; Figure 4a). This switching effect was significant in the comparison between controls and ADHD participants OFF medication (F(1, 35) = 5.62, p = 0.02), with a similar trend for those ON medication

(F(1, 35) = 2.9, p = 0.1). Medication did not significantly affect switching performance (F(1, 35) = 1.8, NS).

The same putative NA mechanism for erratic switching should, according to models and experimental data, also lead to increased reaction time variability (Usher *et al*, 1999; Frank *et al*, in press). Mean RTs did not differ between OFF participants and controls (F(1, 35) = 0.1). Medication actually slowed RTs (F(1, 35) = 15.8, p = 0.003), such that ON patients were significantly slower to respond than controls (F(1, 35) = 4.8, p = 0.03); this slowing may have enabled patients to perform more accurately while on medication (as in speed-accuracy tradeoff effects). To quantify RT variability, we computed the within-subject RT standard deviation in proportion to (ie, normalized by) the individual's mean RT. [This normalization results in a relative measure of variability that is unconfounded by, and controls for, overall differences in RT (since variability necessarily increases with mean). Similar patterns of results were observed with other measures of variability (eg, standard deviations of log-transformed RTs.] Consistent with a common noradrenergic mechanism for variability and erratic switching, these measures were highly correlated in ADHD participants OFF medication (r(17) = 0.71, p = 0.001; Figure 4b), and to a lesser degree but still significantly so in those ON medication (r(18) = 0.47, p = 0.047). In contrast, these measures were not significantly correlated in controls (r(38) = 0.18). Critically, neither RT variability nor switching behavior were correlated with choose-A performance (p's > 0.25), supporting our hypothesized independent DA and NA mechanisms for reinforcement learning and switching/variability deficits in ADHD.

To statistically determine whether the relationship between switching and variability depends on ADHD and medication status, we regressed RT standard deviation against group, medication condition and percent switching, including interactions between these factors and again controlling for mean RT. This analysis revealed that RTs were significantly more variable in OFF participants (normalized variability = 52%) than controls (43%),
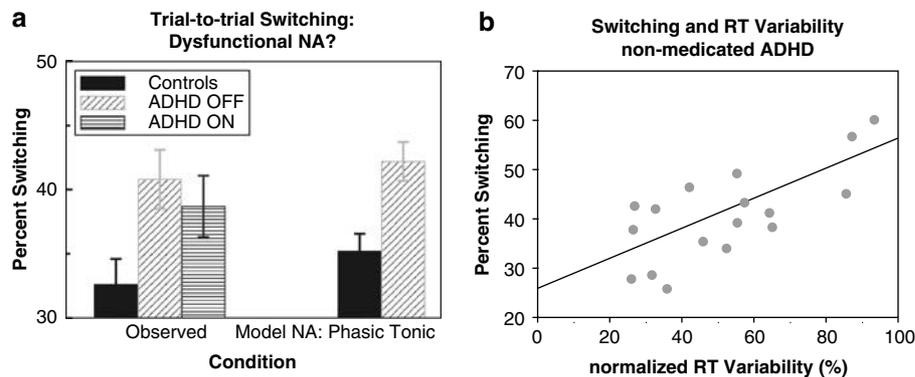

Figure 4 (a) Trial-to-trial response switching (regardless of the previous trial's feedback) in the training phase of the probabilistic learning task. ADHD participants displayed more erratic switching behavior than controls. This effect was captured by tonic versus phasic NA function in our integrated BG/NA model, using the same parameters previously used to simulate reaction time variability (Frank *et al*, in press). Error bars reflect standard error of the mean across 50 networks with different sets of random initial synaptic weights. (b) In nonmedicated ADHD participants, erratic switching behavior was highly correlated (r = 0.71) with variability in reaction times, expressed in terms of standard deviation as a percentage of mean RT. This correlation is consistent with the hypothesized common noradrenergic mechanisms for these effects. Neither of these measures were correlated with positive reinforcement learning performance (see text), consistent with independent DA and NA mechanisms.

(F(1,35) = 5.5, p = 0.02). Medication significantly reduced RT variability (F(1,35) = 5.0, p = 0.03), such that controls and ON participants (normalized variability = 46%) did not differ (F(1,35) = 0.1). The main effect of switching on RT variability was significant (F(1,35) = 5.1, p = 0.03), such that increased switching was associated with increased RT variability. Moreover, this effect was particularly strong in OFF participants, as evidenced by significant interactions with group (ADHD vs control; F(1,35) = 7.3, p = 0.01), and medication status (F(1,35) = 7.1, p = 0.01). Overall these results are consistent with the hypothesis that dysfunctional NA processes in ADHD can lead to both variability and switching behavior.

To explicitly demonstrate that the same high tonic NA mechanism can account for these findings, we ran the same model as in Frank et al (in press) with identical parameters (tonic NA firing rate = 50% maximal) previously used to account for RT variability. Fifty different networks with different sets of random initial synaptic weights were trained to choose among two probabilistically reinforced responses (for simulation details and model equations, see Frank et al (in press)). We then computed the percentage of trials in which the models switched responses according to trial type, regardless of previous feedback, as in the behavioral analysis just reported. Notably, models in the high tonic mode displayed more erratic switching than those in the phasic mode, providing reasonably good quantitative fits to the ADHD data (Figure 4a). Together with our previously reported effects on RT variability, these simulations show that the posited account of common NA mechanisms for these effects is plausible.

Working Memory Task

Our models, together with support from a recent pharmacological study in healthy individuals, predict that the striatal DA mechanisms involved in reinforcement learning may also be important in the working memory domain (Frank and O'Reilly, 2006). More specifically, we have shown that the same Go/NoGo mechanisms that modulate learning and action selection in other frontal regions can also drive the updating of working memory representations in parallel BG-PFC circuits (Frank et al, 2001; Frank, 2005; O'Reilly and Frank, 2006). Go signals cause PFC to update and maintain current sensory information, while NoGo signals prevent updating, enabling robust ongoing maintenance of previously stored information. This account is consistent with striatal activation observed during working memory tasks (Lewis et al, 2004) and with working memory/executive function deficits observed in various patient populations with a BG locus, such as Parkinson's disease (Brown and Marsden, 1990; Owen et al, 1998; Cools et al, 2006). We hypothesized that the reduction in BG DA that led to motivational/reinforcement deficits in ADHD would also cause reduced selective updating and subsequent maintenance of task relevant information.

Although working memory is not always found to be impaired in ADHD (Willcutt et al, 2005; Rhodes et al, 2005), our models predict a more specific impairment in the gating (rather than maintenance per se) of working memory representations. Thus, working memory deficits should be observed when distractors are presented during the delay period, in which case BG gating function is particularly critical for updating only task-relevant information to be maintained. This prediction is consistent with the finding that working memory manipulation—which requires selective updating of some information while working on a subset of a problem—is typically more impaired in ADHD than maintenance (Martinussen et al, 2005); similar updating/manipulation-specific deficits are observed in Parkinson's disease (see Cools, 2006). Finally, it is important to provide a sensitive measure of working memory deficits rather than a more general performance effect that could arise for example from reduced vigilance. Thus, we wanted to test the within-subject tendency to weight internal (working memory) versus external (incoming sensory stimulus) information in guiding responses. Ideally, such a measure is independent of overall performance, as an increased bias to weight working memory representations may not always be beneficial to task performance.

This notion is particularly evident in the 'expectancy' version of the AX-CPT (continuous performance task) working memory task (Servan-Schreiber et al, 1997; Barch et al, 1997, 2001), which we employ here. We also modified the task to include a variable number of distractors during the delay (Frank and O'Reilly, 2006). In the basic task, the participant is presented with sequential letter stimuli (A,X,B,Y; printed in red) and is asked to detect the specific sequence of an A (cue) followed by an X (probe) by pushing the right button (Figure 5). All other cue-probe combinations (A–Y, B–X, B–Y) should be responded to with a left button push. We used both short (1 s) and long (3 s) delays between the cue and probe, with 0–3 distractors presented during the long delay. In the expectancy version used here, the target A–X sequence occurs on 70% of trials, setting up a prepotent expectation of target responses (in contrast to other CPT versions in which targets are presented on a minority of trials); the other sequences are divided equally by the remaining 30% of trials.

This task requires a relatively simple form of working memory where the prior stimulus must be maintained over a delay until the next stimulus appears, so that one can discriminate the target from non-target sequences. Moreover, this task also allows analysis of the type of errors made (Barch et al, 1997; Braver et al, 2001; Frank and O'Reilly, 2006). If participants successfully maintain contextual information (eg, A) in working memory then they will perform well at detecting the A–X target sequence but will likely make more false positive errors on the A–Y sequence (due to anticipation of an X). Context maintenance is particularly critical for performance on the B–X case because one has to maintain the B to know not to respond to the X as a target. The B–Y sequence serves as a control because neither the B or the Y are associated with the target. Furthermore, because the A–X sequence occurs with high (70%) probability, it is not as reliable an indicator of working memory performance because participants can simply learn a prepotent response to stimulus X. Thus, we focus on the B–X and A–Y cases. Specifically, we compute a working memory context index by subtracting percent A–Y accuracy from that of B–X. A positive WM context index indicates greater influence of working memory on choice behavior, whereas a negative context index indicates that
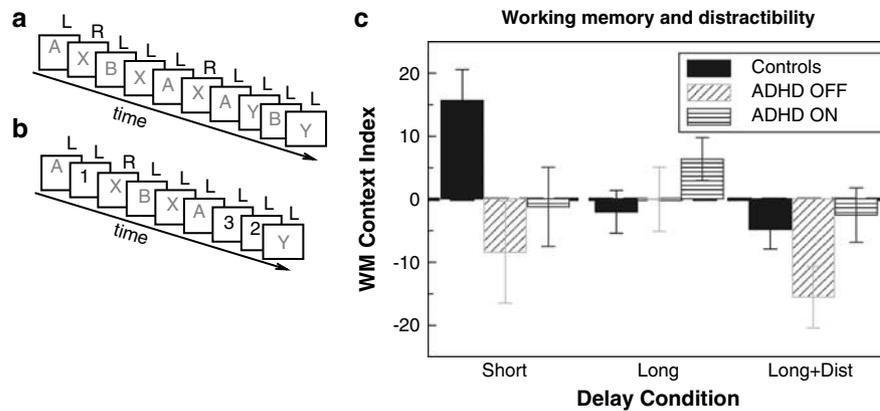
Figure 5 The CPT-A–X task. (a) Short delay, no distractors. Stimuli are presented one at a time in a sequence. The participant responds by pressing the right key (R) to the target sequence, otherwise a left key (L) is pressed. Delay between each stimulus is 1 s. The A–X target sequence occurs on 70% of trials, building up a prepotent expectation for target responses. (b) Variable distractors. Task is the same as in (a) but anywhere from zero to three distractors are presented sequentially during a 3 s delay period. Participants have to respond to distractors with a left button push, but are told to ignore these for the purpose of target detection. In a subsequent attentional-shift, the target sequence consists of previously distracting number stimuli (1–3), and the letter stimuli are now distractors. (c) A–X working memory results for Short (1 s), Long (3 s) delay, and Long + Distractors condition. WM context index = B–X−A–Y, measuring maintenance of context information (A, B) in working memory. As predicted, ADHD was associated with reduced WM context index, which was improved by stimulant medication. This is consistent with increased DA bursting for task-relevant stimuli, reinforcing BG Go signals to update prefrontal cortex working memory representations.

choices are being dictated by incoming stimuli and are not influenced by working memory (Braver *et al*, 2001; Frank and O'Reilly, 2006). [This index can also evaluate the selective contribution of working memory, as opposed to inhibition. That is, both B–X and A–Y cases require suppressing or inhibiting a target response (because both the X in B–X and the A in A–Y are usually associated with targets in the vast majority of A–X sequences). Thus, the context index computes difference between conditions that both involve inhibition, but differ in working memory effects.]

In our BG/PFC framework, we argue that BG Go signals lead to the updating of PFC working memory representations and that DA enhances Go signals, leading to a lowered updating threshold (Frank *et al*, 2001; Frank, 2005; see also Weiner and Joel, 2002; Redgrave *et al*, 1999). Further, DA bursts occur for task-relevant (ie, 'positive') information, making this information more likely to become updated and subsequently maintained (O'Reilly and Frank, 2006; Frank and O'Reilly, 2006). Therefore, we hypothesized that low levels of BG/DA in ADHD would be associated with relatively worse B–X performance and better A–Y performance than controls. By enhancing DA release during bursting, stimulant medication should cause participants to be more likely to update and maintain task-relevant information in working memory. We further predicted that these effects should be particularly evident when distractors are presented during the delay, in which case BG gating signals are necessary to update only task-relevant information and to prevent the updating of distracting information. Predictions for A–X target sequence performance are somewhat more ambiguous, because while working memory gating can increase anticipation and detection of target sequences, good A–X performance could also be accomplished simply by learning a prepotent target response to the X stimulus (which is almost always associated with target responses). Finally, we predicted no effect of ADHD or medication on B–Y performance.

Results and discussion. Figures 5 and 6 show the results for the standard A–X task with short delay, long delay, and long delay with distractors. The patterns observed were consistent with model predictions under various conditions, with particular robustness under distractors: ADHD participants OFF medication performed worst in A–X and B–X trials, which benefit from maintenance of contextual information in working memory. Importantly, these are not global deficits: OFF participants were generally *better* than the other groups at A–Y trials, where contextual maintenance can cause false alarms in anticipation of the prepotent A–X target. Moreover, DA medication reversed all of these effects, improving A–X and B–X performance, while causing more false alarms in A–Y sequences. Finally, B–Y performance, which is not affected by working memory, was similarly good across groups and conditions.

These results were supported by the working memory context index analysis. In the short delay condition, the WM context index was significantly reduced in ADHD participants relative to controls ($F(1, 35) = 8.1$, $p = 0.007$). The within-subject effect of medication in ADHD was not significant ($F(1, 35) = 0.9$). Our inclusion of distractors in some long delay trials was meant to test the role of BG/DA gating of PFC WM representations (Frank and O'Reilly, 2006). Specifically, without gating signals to update and robustly maintain task-relevant information in PFC, distractors would be more likely to interfere with WM representations, leading to a reduced working memory context index. Indeed, WM context index was significantly reduced when distractors were present compared to long delay trials without distractors ($F(1, 35) = 11.1$, $p = 0.002$). The reduced context index under distractors was marginally worse in OFF ADHD participants compared with controls ($F(1, 35) = 2.8$, $p = 0.1$; Figure 5), with no difference between ADHD and controls when no distractors were present ($F(1, 35) = 0.1$). Moreover, medication significantly increased WM context index under distractors ($F(1, 35) = 6.5$, $p = 0.015$), consistent with an increase in
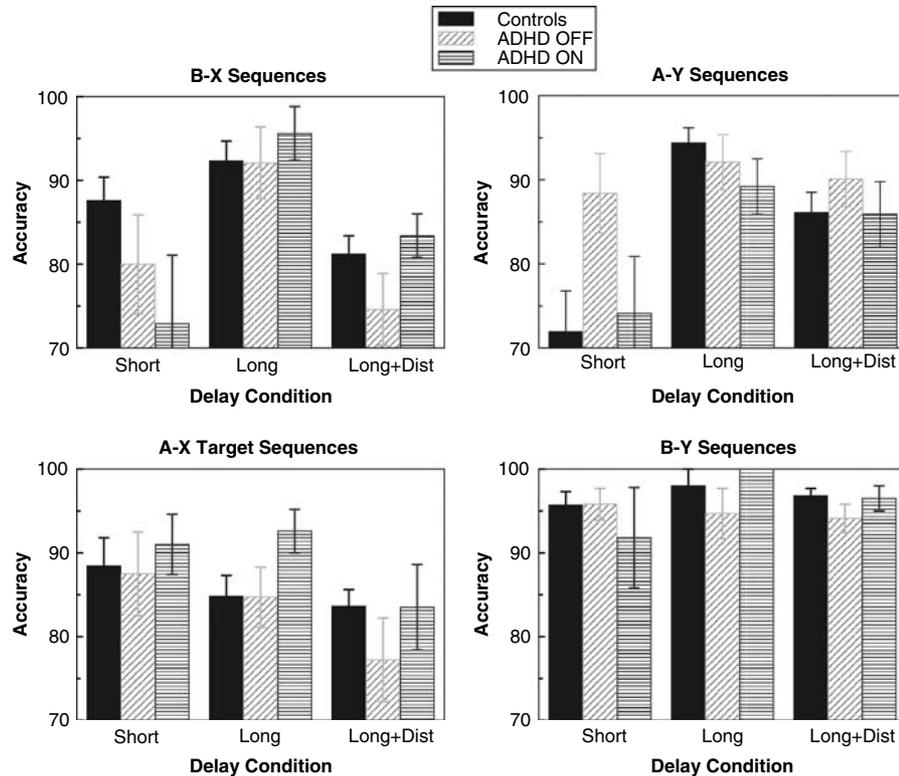
Figure 6 Performance in individual trial types of the working memory task, across delay conditions in ADHD and control participants.

phasic DA signals that support working memory gating of task-relevant information in the face of distractors.

More insight comes from analysis of the individual trial types. First, in the short delay condition, B–X performance did not differ between OFF participants and controls ($F(1,35) = 1.4$, NS), or between medication conditions in ADHD ($F(1,35) = 0.6$). Similarly, there was no effect of ADHD on B–X performance in long delay trials without distractors ($F(1,35) = 0.8$). This lack of B–X effect in the short delay and no distractors suggests that BG/DA gating signals are not as essential for WM processes under these conditions, consistent with the lack of ADHD impairments in fast rate-CPT more generally (Chee *et al*, 1989). However, when distractors were included in long-delay trials, significant B–X deficits were observed in OFF participants relative to controls ($F(1,35) = 4.7$, $p = 0.037$). Critically, these deficits were significantly improved by DA medication ($F(1,35) = 9.0$, $p = 0.005$), such that ON participants' BX performance in the face of distractors did not differ from that of controls ($F(1,35) = 0.2$).

For A–Y, as noted above, performance can suffer from enhanced working memory context maintenance. Indeed, A–Y performance was significantly *better* in OFF ADHD participants than controls in short delay trials ($F(1,35) = 4.5$, $p = 0.04$). Medication marginally worsened performance and caused more false alarms ($F(1,35) = 2.9$, $p = 0.09$), such that ON participants' A–Y performance did not differ from that of controls ($F(1,35) = 0.1$). Thus, although gating signals may not be necessary for context maintenance under these short delay conditions, enhanced Go signals (in healthy conditions, or via DA medication in ADHD) can nevertheless increase anticipation of A–X target sequences and therefore cause A–Y false alarms.

Consistent with this depiction, medication improved A–X target detection in both short delay ($F(1,35) = 3.8$, $p = 0.06$), and long delay trials ($F(1,35) = 8.3$, $p = 0.007$). A–X target performance was not impaired in ADHD participants relative to controls in either short ($F(1,35) = 1.0$) or long delay trials ($F(1,35) = 0.7$). These results again support the notion that working memory gating is not as essential for performing well on prepotent A–X target sequences but that enhanced BG/DA gating signals can nevertheless cause more preparatory anticipation and therefore better performance. Finally, there was no difference between ADHD participants and controls in B–Y performance ($F(1,35) = 0.04$) and no effect of medication ($F(1,35) = 0.66$).

Relationship between Reinforcement Learning and Working Memory in ADHD

As described above, our models predict that the same mechanism of reduced striatal DA, acting on different fronto-striatal circuits, leads to ADHD-related deficits in both positive reinforcement learning and working memory gating. Our results were consistent with this notion: both deficits were observed and were ameliorated by stimulant medications that increase DA levels. If deficits on both tasks arise from a common mechanism, a stronger prediction is that performance on the two measures should be correlated and similarly affected by DA manipulation. First, we found that in nonmedicated participants, positive reinforcement 'Go' learning (choose-A) performance was correlated with performance in B–X working memory sequences under distractors ($r(18) = 0.52$, $p = 0.027$), but not when no distractors were present ($r(18) = -0.08$). This is consistent

with the notion that striatal Go signals are necessary for gating of information to be robustly maintained in the face of ongoing interference (Frank *et al*, 2001), as measured by the B–X distractor condition, and that selective gating is less critical when no distractors are present. Furthermore, striatal NoGo signals are also important in the models for preventing the gating of distractors into working memory (O'Reilly and Frank, 2006), such that distractor performance may also correlate with impaired NoGo learning in ADHD. Indeed, aversive 'NoGo' reinforcement learning (avoid-B) performance was correlated with accuracy in the B–X distractor condition ($r(18) = 0.59$, $p = 0.01$), and not when no distractors were present ($r(18) = -0.26$, NS). Overall, these analyses support the notion that both Go and NoGo mechanisms are critical for updating task-relevant information and preventing it from subsequent interference.

More specific predictions come from analysis of medication effects on learning and working memory. We hypothesized that the mechanism by which medication improves working memory is via increases in phasic DA and associated striatal Go signals. This hypothesis predicts that the extent to which medications selectively improved positive reinforcement learning (depending on phasic DA) should be predictive of their improvement working memory gating. We therefore computed the correlation between medication improvements in Go relative to NoGo reinforcement learning (choose-A relative to avoid-B; this difference should reflect improvements due to phasic DA increases) and their improvement in B–X trials with distractors. Indeed, relative Go-NoGo improvements were significantly correlated with B–X accuracy improvements under distractors ($r(15) = 0.59$, $p = 0.02$; Figure 7), and not in B–X trials without distractors ($r(15) = -0.13$, NS). Further, improvements in A–X performance under distractors were not correlated with relative Go reinforcement learning improvements ($r(15) = -0.12$, NS), again supporting the notion that good A–X performance can result simply from a prepotent bias to respond to the X, and is not critically dependent on working memory gating. Finally, there were also no significant correlations between improvements in B–X and those in putative noradrenergic effects of reaction time variability or switching behavior (p's > 0.1 and 0.4, respectively), supporting our previous analysis that positive Go learning performance also did not correlate with these measures (p's > 0.25).

GENERAL DISCUSSION

Taken together, our findings support a unified neurocomputational account of cognitive and motivational deficits in ADHD. First, ADHD participants were impaired at positive reinforcement learning (in both training/test measures) and selective working memory gating, supporting our models' suggestion that both of these processes depend on dopamine signals within the basal ganglia (Frank, 2005; O'Reilly and Frank, 2006), which are reduced in ADHD. Second, these impairments were reversed by stimulant medications which enhance DA levels and associated striatal Go signals (eg, Yano and Steiner, 2005), particularly in conditions that depend on phasic DA in our models.
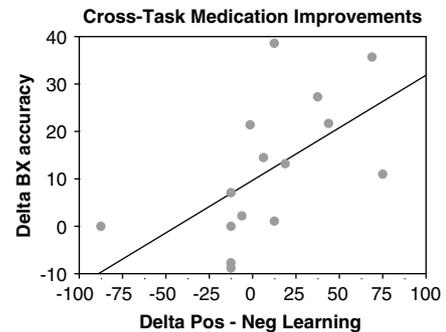

Figure 7 Medication improvements in positive relative to negative reinforcement were correlated with medication improvements in B–X performance in the distractor condition of the working memory task ($r(15) = 0.59$, $p = 0.02$). According to the BG-PFC models, this condition is particularly sensitive to increases in phasic DA, because one has to update the B and maintain it in the face of distractors in order to know how to respond to the X probe.

These include positive reinforcement learning and B–X performance under distracting conditions in the working memory task. Critically, the extent to which medications improved Go reinforcement learning was correlated with their efficacy in improving working memory gating, suggestive of common DA mechanisms. Third, ADHD participants also showed erratic behavioral switching from trial to trial, which was correlated with increased reaction time variability. Both of these effects are predicted by neurocomputational models of NA dysfunction (Aston-Jones and Cohen, 2005; McClure *et al*, 2006; Frank *et al*, in press) and were also observed when selectively manipulating NA in our model (Figure 4). Finally, supporting our hypothesis for independent mechanistic sources of DA and NA dysfunction, putative DA deficits were uncorrelated with putative NA deficits.

In the discussion that follows, we address the DA and NA hypotheses in turn, before commenting on the combined theory and relationship to other theories of ADHD.

Basal Ganglia/Dopamine

As reviewed above, various lines of evidence point to DA dysfunction in the basal ganglia of ADHD participants. By virtue of interactions with multiple frontal circuits, it is possible that this single low-level mechanism may be responsible for diverse behavioral effects at the systems level. Specifically, reduced striatal DA signals would decrease Go signals for reinforcing appropriate motor behaviors represented in pre/motor regions (Frank, 2005). Similarly, by interacting with prefrontal regions, Go signals are particularly important for selectively updating task-relevant information to be robustly maintained, especially in the face of ongoing distractors (Frank *et al*, 2001; O'Reilly and Frank, 2006). [Without BG Go signals, cortico-cortical projections can still allow a sensory stimulus to reach and activate PFC. However, without the gating signal this information is less robustly maintained, and will be more susceptible to interfering effects of any subsequent distractor.] Thus, reduced striatal Go signals may lead to apparent hypofrontality due to reductions in the selective maintenance of task-relevant relative to distracting

information. Further, the same effects may apply with respect to ventral striatum and the updating of orbitofrontal working memory representations of reward value (see Frank and Claus, 2006). In this case, DA reductions would lead to impairments in the updating and subsequent maintenance of large magnitude, long-term reward values to bias behavior and motivational processes. This account is consistent with a steeper and shorter delay-of-reinforcement gradient in ADHD (Sagvolden et al, 1998), and with observations that reinforcement motivation needs to be particularly high in ADHD participants (Luman et al, 2005). More generally, our hypothesized common mechanism for reinforcement and working memory deficits from a modeling perspective is consistent with Sagvolden et al's (2005) suggestion that working memory and delay of reinforcement gradient effects in ADHD are linked through reduced dopaminergic efficacy. Support for our particular mechanism comes from a recent study showing that stimulants induce striatal Go activity and learning, which is highly correlated with enhanced frontal cortical gene expression in specific cortico-striatal circuits (Yano and Steiner, 2005).

There may be alternative explanations for working memory deficits in ADHD. For example, it is possible that ADHD participants showed excessive, rather than reduced, updating. In this case, participants would be more likely to update distractor stimuli, which could lead to distractibility by interfering with previously updated information. This seems unlikely, particularly given that reduced WM was observed when no distractors were presented in the short delay condition. In contrast, other pharmacological data are more consistent with this overall increased BG updating/distractibility in the same tasks. Specifically, we tested healthy participants taking the drug cabergoline, a D2 agonist that should cause excessive updating (by inhibiting the NoGo pathway and lowering the threshold for Go responses; Frank and O'Reilly, 2006). In that case, participants also showed impaired WM performance under distractors, but *better* performance in the short delay zero-distractor condition—in which case enhanced BG updating/facilitation of working memory representations can only help. Thus, our finding here that ADHD participants showed reduced WM context index in both short delay and distractor conditions is more consistent with reduced BG Go signals for gating WM representations.

A second alternative is that maintenance itself, rather than updating, is impaired in ADHD. However, a maintenance deficit would predict that WM performance is increasingly impaired with increasing delay (even in the absence of distractors). This was clearly not the case in our data: if anything WM performance was better in the long delay without distractors than in the short delay case. Thus, the most parsimonious account of our data, and consistent with the reinforcement learning data, is that ADHD is associated with reduced BG Go signals for gating information to be maintained in PFC.

Comparison with parkinson's disease. One problem with the hypodopaminergic hypothesis usually unaddressed is why are not low DA levels in ADHD associated with Parkinson-like symptoms? First, it is likely that tonic DA levels are much lower in Parkinson's, given that symptoms do not arise until DA is depleted by approximately 75–80%. Second, whereas Parkinson's patients simply do not have DA available, DA synthesis and availability should be intact in ADHD. Individuals with ADHD may try to self-regulate and increase their DA levels, as seen in rats who self administer more amphetamine when DA receptors are partially blocked pharmacologically (Robbins and Everitt, 1999). Intriguingly, ADHD participants may achieve these DA increases by their own hyperactive movements (Frank et al, in press): matrix neurons in the dorsal striatum involved in motor selection may disinhibit DA release via striatonigral projections (eg, Joel and Weiner, 2000). These motor-induced increases in DA levels would not be detected by neuroimaging studies, which require subjects to remain still in the scanner; nevertheless, the hypothesis is supported by observations that more hyperactive children have higher striatal D2 receptor binding (Jucaite et al, 2005), which may reflect low tonic DA during rest. More generally, this account is consistent with theories positing that hyperactivity is a compensatory process for the lack of external stimulation (eg, Barkley, 1997; Sonuga-Barke, 2003).

Another question that arises from our reinforcement learning data is why negative feedback (avoid-B) performance was impaired in OFF ADHD participants relative to controls, given their low levels of striatal DA (which should facilitate NoGo learning). Indeed, Parkinson's patients actually showed somewhat *better* NoGo learning (Frank et al, 2004). This discrepancy is potentially explained by (i) far lower DA levels and (ii) supersensitivity of D2 receptors in Parkinson's disease (eg, Rinne et al, 1990). This combination would make striatal neurons in PD particularly sensitive to DA dips for learning to avoid negative outcomes, even in the presence of tonically low DA levels (for discussion see Frank and O'Reilly, 2006). Further, we have argued that the most relevant comparison for testing specific predictions is the within subject effect of controlled manipulations, because other factors can contribute to overall performance differences in patient groups (Frank et al, 2004). In this study, we found that stimulant medication enhanced Go but not NoGo learning in ADHD. Moreover, trial-to-trial analysis during training revealed selective ADHD deficits in positive reinforcement learning, consistent with reduced DA levels.

Noradrenaline

Although DA dysfunction in ADHD is well supported, growing evidence also points to NA dysfunction (Russell and Wiggins, 2000; Russell et al, 2000; Swanson et al, 2006; Michelson et al, 2001; Overtoom et al, 2003). In particular, standard stimulants such as amphetamine and methylphenidate increase both DA and NA levels by acting on reuptake transporters for both neurotransmitters (Madras et al, 2005; Berridge et al, 2006). Interestingly, whereas the dopamine transporter density is very high in striatum and very low in prefrontal cortex, the opposite relationship is observed for the noradrenaline transporter (eg, Madras et al, 2005). Thus, it is likely that therapeutic effects of NA manipulation in ADHD are primarily mediated in frontal cortex. This possibility fits nicely with computational models and associated neurophysiological experiments

which show that phasic NA release in frontal cortex is associated with periods of focused attention, infrequent target detection, and good task performance (Aston-Jones and Cohen, 2005). These authors simulated these effects by having NA modulate the gain of the activation function in cortical response units (Usher *et al*, 1999), showing that phasic NA release leads to 'sharper' cortical representations and a tighter distribution of reaction times. In contrast, a relatively high tonic (but low phasic) NA state during periods of poor task performance was associated with more RT variability in both their models and monkey behavioral data. They further hypothesized that this NA state during poor performance may be adaptive, in that it may enable the representation of alternate competing cortical actions during exploration of new behaviors.

As previously noted, this model has clear implications for NA dysfunction in ADHD (Frank *et al*, in press). In particular, a high tonic/low phasic NA state in ADHD would lead to noisier frontal cortical representations, which could cause variability in reaction times, via effects in pre/motor areas. Indeed, various studies show that ADHD participants show increased within-subject reaction time variability (Leth-Steensen *et al*, 2000; Lijffijt *et al*, 2005; Castellanos *et al*, 2005). Interestingly, the same NA mechanism also predicts that noise in frontal motor representations may cause erratic exploratory behavior, or in other words, reduced consistency of choice responses. In the present study we found compelling support for this account: not only were ADHD participants more likely to display erratic switching in their behavioral choices from trial to trial, but increased switching behavior was correlated with increased variability in reaction times, supporting a common NA source for these behavioral effects and captured in our model.

Our NA account contrasts with a tentative proposal that RT variability in ADHD stems from DA mechanisms that cause low frequency (multisecond) neural oscillations in BG circuits (Castellanos *et al*, 2005). However, in the rat studies on which that theory is based, slow BG oscillations depended on an *increase* in DA (Ruskin *et al*, 2003), whereas ADHD is associated with reduced DA. Further, in primates, slow BG oscillations were unaffected by DA manipulation (Wichmann *et al*, 2002). Other accounts suggest that variability in ADHD could arise from inefficient learning, due to slower ability to develop predictable behavior and unlearn ineffective responses (Sagvolden *et al*, 2005). In our models, poor Go learning associated with reduced DA might indirectly lead to RT variability: in the absence of efficacious BG Go signals, noisy cortical activity could have a greater effect on network behavior. Nevertheless, it is this noisy activity itself that is thought to be enhanced under conditions of high tonic NA. In effect, we suggest that reduced BG/DA signals do not themselves cause variability but could nevertheless potentially 'reveal' the more direct effects of NA on cortical noise and associated variability. [Indeed, the high tonic NA simulated in the models to produce RT variability was carried out *after* learning had been successfully achieved and was tested across 5000 trials during which no further learning occurred (Frank *et al*, in press). This demonstrates that high tonic NA is sufficient in and of itself to account for variability, without requiring dynamic interactions with trial to trial

learning effects.] This more direct effect of NA on variability is consistent with data from a recent study measuring urinary excretion of catecholamine metabolites in children with ADHD, which showed that RT variability correlated with NA, but not DA, metabolites (Llorente *et al*, 2006).

To more directly test whether DA depletion in humans could account for RT variability in the same task reported here, we analyzed data from PD patients from Frank *et al* (2004). These data provide an opportunity to test a key prediction: whereas DA levels are reduced in both PD and ADHD, tonic NA levels are thought to be high in ADHD but are depleted in PD (Hornykiewicz and Kish, 1987; Bertrand *et al*, 1997). Thus, if RT variability stems from reduced DA levels, then this variability should also be seen in PD. However, if high tonic NA is the primary source of variability, then if anything variability should be *reduced* in PD patients, due to low NA levels. Indeed, nonmedicated PD patients showed somewhat reduced normalized RT variability compared with controls ($F(1, 26) = 2.7$, $p = 0.11$), an effect that was significant when compared with medicated PD patients ($F(1, 26) = 4.7$, $p = 0.04$). This medication effect can be explained by observations that PD-related medication enhance NA levels (eg, Stryjer *et al*, 2005), thereby counteracting NA depletion in PD, and increasing RT variability. Taken together, ours and other results, including metabolite studies in human ADHD participants and direct LC/NA recordings in monkeys (Llorente *et al*, 2006; Usher *et al*, 1999), suggest that RT variability in ADHD stems from high tonic NA levels and is unlikely to be accounted for by DA depletion.

Note that although there may be intrinsic NA dysfunction in ADHD, this need not be the case. It is plausible that a high tonic/low phasic NA state is encountered as a secondary reaction to deficits caused by other related mechanisms, such as reduced striatal facilitation of responses, reduced grey matter, disorganization of representations, and associated poor task performance. Similarly, increases in NA by atomoxetine could compensate for these other deficits.

Relationship to Theories on the Independence of Cognitive and Motivational Deficits in ADHD

Others have proposed theories to explain the independence of cognitive and motivational deficits in ADHD, based on behavioral evidence showing that response inhibition deficits are independent from motivational deficits (Solanto *et al*, 2001). In particular, Sonuga-Barke (2003) and Castellanos *et al* (2006) propose a dual-pathway model, which implicates separate cortico-striatal circuits in motivational ('hot') and executive ('cool') dysfunction in ADHD. Specifically, response inhibition deficits are proposed to reflect damage to prefrontal-dorsal striatal circuits, whereas motivational and reward deficits stem from dysfunction in ventral striatal-orbitofrontal (OFC) circuits.

Our theory has important similarities and differences with that of Castellanos *et al* (2006). First, central to both accounts is a core motivational deficit in ADHD. Second, our models are broadly consistent with the proposed neurobiological underpinnings: we have explicitly simulated the contributions of the striatum, orbitofrontal cortex and dopamine to motivation, reward learning, and decision

making (Frank and Claus, 2006). This model suggests that reduced DA in striato-orbitofrontal circuits would lead to impaired representation of large rewards, consistent with evidence that orbitofrontal areas are involved in the selection of large delayed rewards over smaller but more immediate rewards (Mobini *et al*, 2002; McClure *et al*, 2004). Disruption of this circuitry is supported by evidence that ADHD participants are often found to have an unusually strong preference for small immediate rewards over larger delayed rewards (Sonuga-Barke, 2005; but see Scheres *et al*, 2006a).

Despite the commonalities among our models, we maintain that striatal DA reduction in ADHD may be a common source of both these 'hot' motivational deficits as well as the 'cool' working memory deficits (see also Sagvolden *et al*, 2005). Increased DA transporter expression in ADHD (Dougherty *et al*, 1999; Krause *et al*, 2000) would presumably lead to non-selectively reduced DA across the entire striatum, leading to impairments in functions depending on multiple cortico-striatal circuits. We suggest that a more likely source of symptom independence is a completely different mechanism than DA, and propose NA to be one such mechanism. Of course other independent mechanisms are possible, such as the consistently reduced grey matter in the cerebellum (see Krain and Castellanos, 2006). Even within the DA system, differential dopaminergic genes have been shown to affect prefrontal versus striatal volume (Durston *et al*, 2005), and these could also lead to dissociable symptoms.

An obvious prediction of our theory is that if one were to independently manipulate DA and NA levels in ADHD, dissociations could be observed for the two kinds of behavioral effects. Presently, this may be difficult: most stimulant medications, including the variants taken by participants in this study, elevate both DA and NA levels (Madras *et al*, 2005). While the newer drug atomoxetine acts preferentially on the NA transporter (Michelson *et al*, 2001) and increases frontal NA levels, it also increases PFC DA levels (via indirect effects on the NA transporter; Bymaster *et al*, 2002; Madras *et al*, 2005). Nevertheless, this drug does not elevate DA levels in striatum, where the NA transporter is scant (Bymaster *et al*, 2002; Madras *et al*, 2005). Thus one might predict that in contrast to other medications, atomoxetine should not selectively improve positive relative to negative reinforcement learning. Rather, by modulating PFC DA, the drug could improve vigilance and lead to better performance overall. In addition, by modulating PFC NA, the drug should have more reliable effects on RT variability and erratic switching behavior.

Finally, it is plausible that NA cortical effects are involved in response inhibition deficits (in addition to variability and erratic behavior), whereas BG/DA effects are involved in motivational/reward processes, supporting the independence of these symptoms (Solanto *et al*, 2001). This notion is consistent with recent evidence that atomoxetine improves response inhibition (eg, Swanson *et al*, 2006; Michelson *et al*, 2001; Overtoom *et al*, 2003; Chamberlain *et al*, 2006). However, the current evidence does not rule out a role for DA in inhibition, given the beneficial effects of traditional stimulants on inhibition, which could act via DA, NA or a combination of the two. Evidence for a BG role in inhibition is suggested by observations that children with ADHD show reduced striatal activation when responses are to-be-inhibited (eg, Durston *et al*, 2003). We are currently investigating the interactions of NA effects in premotor cortical regions with BG/DA effects in our models. Ultimately, a robust model will investigate interactions between these various systems.

Conclusions

In summary, our results demonstrate the usefulness of neurocomputational approaches for hypothesis generation and testing in ADHD, which may have both practical and theoretical implications. For example, if replicated in younger children, our reinforcement learning results may have implications for motivational strategies in parents and teachers of medicated ADHD participants. Our findings and models suggest that these participants may respond much better to reward-based motivations then to punishments. Theoretically, our framework suggests a common dopaminergic mechanism for reinforcement learning and working memory deficits and an independent noradrenergic mechanism for reaction time variability and erratic behavior. We look forward to future genetic, neuroimaging, and pharmacological research with behavioral paradigms such as those used in the present study to test these ideas.

METHODS

Procedures were approved by the University of Colorado Human Research Committee.

Participants

Eighteen ADHD participants and 21 controls who did not differ in age or SAT scores (p's > 0.25). were included. ADHD participants (11 men, 7 women) ranged in age from 18 to 22 ($M = 20$, $s = 1.49$), with a mean SAT score of 1194. All participants with ADHD had a current prescription for stimulant medication to treat ADHD symptoms, as prescribed by their physician. These medications included methylphenidate ($N = 3$), pemoline ($N = 1$), amphetamine ($N = 1$), and an amphetamine–dextroamphetamine combination ($N = 13$). All participants completed two sessions and were paid 30 dollars for each session. ADHD participants were tested in two sessions, once on medication and once off medication, with the order counterbalanced across participants. For the off-medication session, participants abstained from taking their medication for a minimum of 24 h before the session. Twenty one control participants volunteered by signing up on a website advertising paid experiments at the University of Colorado at Boulder, and completed two sessions, not on medication for either session. Control participants completed a screening measure on which they rated each of the nine DSM-IV inattention symptoms and nine DSM-IV hyperactivity-impulsivity symptoms on a 0–3 scale (Barkley and Murphy, 1998). We excluded participants who rated at least six symptoms from either category (inattentive or hyperactive) 2 or higher. Two participants were excluded using these criteria. The remaining 19 control participants (8 men, 11 women) ranged in age from 18 to 22 years ($M = 20$, $s = 1.43$) and had a mean SAT score of 1225.

In addition to the tasks described here, participants also completed other tasks to address questions about the relationship between time perception and working memory in ADHD, to be reported later (Santamaria *et al*, in preparation). The order of these tasks was counterbalanced across participants.

Reinforcement Learning Task

Procedures for this task have been described elsewhere (Frank *et al*, 2004; Frank and O'Reilly, 2006). Briefly, participants sit in front of a computer screen in a lighted room and view pairs of visual stimuli that are not easily verbalized (Japanese Hiragana characters, see Figure 2). These stimuli are presented in black on a white background, in 72 pt font. They press keys on the left or right side of the keyboard depending on which stimulus they choose to be 'Correct'. Note that the forced-choice nature of the task controls for any differences in overall motor responding. Visual feedback is provided (duration 1.5 s) following each choice (the word 'Correct!' printed in blue or 'Incorrect' printed in red). If no response is made within 4 s, the words 'no response detected' are printed in red.

We enforced a performance criterion (evaluated after each training block of 60 trials) in an attempt to equate training performance at the time of test. Because of the different probabilistic structure of each stimulus pair, we used a different criterion for each (65% A in AB, 60% C in CD, 50% E in EF). (In the EF pair, stimulus E is correct 60% of the time, but this is particularly difficult to learn. We therefore used a 50% criterion for this pair simply to ensure that if participants happened to 'like' stimulus F at the outset, they nevertheless had to learn that this bias was not going to consistently work.) The participant advanced to the test session if all these criteria were met, or after six blocks (360 trials) of training. The test session involved presenting the same training pairs in addition to all novel combinations of stimuli, in random sequence. They were instructed (prior to the test phase) to use 'gut instinct' if they did not know how to respond to these novel pairs. Each test pair was presented six times for a maximum of four seconds duration, and no feedback was provided.

Working Memory Procedures

The participant is presented with sequential letter stimuli (A,X,B,Y) printed in red, and is asked to detect the specific sequence of an A (cue) followed by an X (probe) by pushing the right button. All other cue-probe combinations (A–Y, B–X, B–Y) should be responded to with a left button push. The target A–X sequence occurs on 70% of trials, and the other sequences are divided equally by the remaining 30% of trials. The working memory segment begins with the standard task with no distractors and an inter-stimulus interval of one second. Each stimulus is presented for 500 ms. There are 50 sequences in the standard task.

Next, subjects were told to continue looking for the A–X target sequence, and to ignore any distractor stimuli (single white digits (1,2,3,4)) that may appear during a 3-s cue-probe delay interval. Each stimulus is presented for 500 ms, and the delay between stimuli and probes is 3 s. Anywhere from 0 to 3 distractors are presented during the delay

period for 333 ms each. The distractors are spaced out evenly throughout the 3 s delay period. When one distractor is presented, there is a 1333-ms delay between the cue and the distractor and between the distractor and the target. For two distractors, the delay between each item is 778 ms. For three distractors, it is 500 ms. Participants have to respond to each distractor with a left button push to ensure that they encode them, but are told to ignore them for the purpose of target detection (Braver *et al*, 2001; Frank and O'Reilly, 2006). In this long delay/distractor segment, there are 128 trials of which 80 are target sequences.

All results are presented across all trials. In actuality, subjects were instructed to respond to a different target sequence in different blocks of the experiment (Frank and O'Reilly, 2006). For simplicity, 'A–X' performance refers to performance on target sequences, 'A–Y' performance refers to performance on non-targets that share the first stimulus cue with the target, etc. Note that the results do not depend on this averaging across blocks (except for the greater number of trials); all the patterns reported held in just the first block in which targets were always A–X.

Data Analysis

In order to be consistent across all data analyses, we performed the same statistical test for each analysis. We used SAS v8.0 PROC MIXED to examine both between and within subject differences, using unstructured covariance matrices (which does not make any strong assumptions about the variance and correlation of the data, as do structured covariances). In all analyses, we controlled for session number (Frank and O'Reilly, 2006). In the procedural learning tasks, another factor of positive/negative test condition was included, whereas in the working memory segments, a distractor factor was included. In the reaction time variability analyses, we included a switching factor term and its interactions with group and medication status, to determine whether the coupling between switching and variability depended on ADHD and medication status. Where indicated, we tested for specific planned contrasts. In these contrasts, the number of degrees of freedom reflects the entire sample, and not just the participants involved in the particular contrast, because the mixed procedure analyzes both between and within effects, and controls for other variables of interest (eg, session) that apply across all participants. This procedure uses all of the data to provide a more stable estimate of the error term.

REFERENCES

Aase H, Meyer A, Sagvolden T (2006). Moment-to-moment dynamics of ADHD behaviour in South African children. *Behav Brain Funct* **2**: 11.

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

APA (ed) (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Press: Washington, DC.

Arnsten AF, Steere JC, Hunt RD (1996). The contribution of alpha 2-noradrenergic mechanisms of prefrontal cortical cognitive function. Potential significance for attention-deficit hyperactivity disorder. *Arch Gen Psychiatry* 53: 448–455.

Aron AR, Poldrack RA (2005). The cognitive neuroscience of response inhibition: relevance for genetic research in attention-deficit/hyperactivity disorder. *Biol Psychiatry* 57: 1285–1292.

Aston-Jones G, Cohen JD (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annu Rev Neurosci* 28: 403–450.

Barch DM, Braver TS, Nystom LE, Forman SD, Noll DC, Cohen JD (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia* 35: 1373–1380.

Barch DM, Carter CS, Braver TS, Ans Sabb FW, MacDonald AR, Noll DC et al (2001). Selective deficits in prefrontal cortex function in medication-naive patients with schizophrenia. *Arch Gen Psychiatry* 58: 280–288.

Barkley RA (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psycholo Bull* 121: 65–94.

Barkley RA, Murphy K (1998). *Attention-Deficit Hyperactivity Disorder: A Clinical Workbook*. Guilford Press: New York.

Berridge CW, Devilbiss DM, Andrzejewski ME, Arnsten AFT, Kelley AE, Schmeichel B et al (2006). Methylphenidate preferentially increases catecholamine neurotransmission within the prefrontal cortex at low doses that enhance cognitive function. *Biol Psychiatry* 60: 1111–1120.

Bertrand E, Lechowicz W, Szpak GM, Dymecki J (1997). Qualitative and quantitative analysis of locus coeruleus neurons in Parkinson's disease. *Folia Neuropathol* 35: 80–86.

Biederman J, Faraone SV (2002). Current concepts on the neurobiology of attention-deficit/hyperactivity disorder. *J Attention Disorders* 6(Suppl 1): S7–S16.

Braver TS, Barch DM, Keys BA, Carter CS, Cohen JD, Kaye JA et al (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. *J Exp Psychol Gene* 130: 746–763.

Brown RG, Marsden CD (1990). Cognitive function in parkinson's disease: From description to theory. *Trends Neurosci* 13: 21–29.

Bymaster FP, Katner JS, Nelson DL, Hemrick-Luecke SK, Threlkeld PG, Heiligenstein JH et al (2002). Atomoxetine increases extracellular levels of norepinephrine and dopamine in prefrontal cortex of rat: a potential mechanism for efficacy in attention deficit/hyperactivity disorder. *Neuropsychopharmacology* 27: 699–711.

Castellanos FX, Sonuga-Barke EJS, Milham MP, Tannock R (2006). Characterizing cognition in ADHD: beyond executive dysfunction. *Trends Cogn Sci* 10: 117–123.

Castellanos FX, Sonuga-Barke EJS, Scheres A, Di Martino A, Hyde C, Walters JR (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biol Psychiatry* 57: 1416–1423.

Chamberlain SR, Muller U, Blackwell AD, Clark L, Robbins TW, Sahakian BJ (2006). Neurochemical Modulation of Response Inhibition and Probabilistic Learning in Humans. *Science* 311: 861–863.

Chee P, Logan G, Schachar R, Lindsay P, Wachsmuth R (1989). Effects of event rate and display time on sustained attention in hyperactive, normal, and control children. *J Abnorm Child Psychol* 17: 371–391.

Cools R (2006). Dopaminergic modulation of cognitive function—implications for ʟ-DOPA treatment in Parkinson's disease. *Neurosci Biobehav Rev* 30: 1–23.

Cools R, Barker RA, Sahakian BJ, Robbins TW (2001). Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cerebral Cortex* 11: 1136–1143.

Cragg SJ, Hille CJ, Greenfield SA (2002). Functional domains in dorsal striatum of the nonhuman primate are defined by the dynamic behavior of dopamine. *J Neurosci* 22: 5705–5712.

Diamond A (2005). Attention-deficit disorder (attention-deficit/hyperactivity disorder without hyperactivity): a neurobiologically and behaviorally distinct disorder from attention-deficit/hyperactivity disorder (with hyperactivity). *Dev Psychopathol* 17: 807–825.

Dougherty DD, Bonab AA, Spencer TJ, Rauch SL, Madras BK, Fischman AJ (1999). Dopamine transporter density in patients with attention deficit hyperactivity disorder. *Lancet* 354: 2132–2133.

Durston S, Fossella JA, Casey BJ, Pol HEH, Galvan A, Schnack HG et al (2005). Differential effects of DRD4 and DAT1 genotype on fronto-striatal gray matter volumes in a sample of subjects with attention deficit hyperactivity disorder, their unaffected siblings, and controls. *Mol Psychiatry* 10: 678–685.

Durston S, Tottenham NT, Thomas KM, Davidson MC, Eigsti I-M, Yang Y et al (2003). Differential patterns of striatal activation in young children with and without ADHD. *Biol Psychiatry* 53: 871–878.

Faraone SV, Perlis RH, Doyle AE, Smoller JW, Goralnick JJ, Holmgren MA et al (2005). Molecular genetics of attention-deficit/hyperactivity disorder. *Biol Psychiatry* 57: 1313–1323.

Frank MJ (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. *J Cogn Neurosci* 17: 51–72.

Frank MJ (2006). Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Networks* 19: 1120–1136.

Frank MJ, Claus ED (2006). Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev* 113: 300–326.

Frank MJ, O'Reilly RC (2006). A mechanistic account of striatal dopamine function in human cognition: Psychopharmacological studies with cabergoline and haloperidol. *Behav Neurosci* 120: 497–517.

Frank MJ, Loughry B, O'Reilly RC (2001). Interactions between the frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, and Behavioral Neuroscience* 1: 137–160.

Frank MJ, Scheres A, Sherman SJ (in press). Understanding decision making deficits in neurological conditions: Insights from models of natural action selection. *Philos Trans Roy Soc-B*.

Frank MJ, Seeberger LC, O'Reilly RC (2004). By carrot or by stick: cognitive reinforcement learning in Parkinsonism. *Science* 306: 1940–1943.

Frank MJ, Woroch BS, Curran T (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron* 47: 495–501.

Grace AA (2001). Psychostimulant actions on dopamine and limbic system function: relevance to the pathophysiology and treatment of ADHD. In: Solanto MV and Arnsten AF (eds). *Stimulant Drugs and ADHD: Basic and Clinical Neuroscience*. Oxford University Press: New York. pp 134–155.

Hornykiewicz O, Kish SJ (1987). Biochemical pathophysiology of Parkinson's disease. *Adv Neurol* 45: 19–34.

Joel D, Weiner I (2000). The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience* 96: 451.

Jucaite A, Fernell E, Halldin C, Forssberg H, Farde L (2005). Reduced midbrain dopamine transporter binding in male adolescents with attention-deficit/hyperactivity disorder: association between striatal dopamine markers and motor hyperactivity. *Biol Psychiatry* **57**: 229–238.

Krain AL, Castellanos FX (2006). Brain development and ADHD. *Clin Psychol Rev* **26**: 433–444.

Krause K, Dresel SH, Krause J, Kung HF, Tatsch K (2000). Increased striatal dopamine transporter in adult patients with attention deficit hyperactivity disroder: effects of methylphenidate as measured by single photon emission computed tomography. *Neurosci Lett* **285**: 107–110.

Leth-Steensen C, Elbaz ZK, Douglas VI (2000). Mean response times, variability, and skew in the responding of adhd children: a response time distributional approach. *Acta Psychol* **104**: 167–190.

Lewis SJ, Dove A, Robbins TW, Barker RA, Owen AM (2004). Striatal contributions to working memory: a functional magnetic resonance imaging study in humans. *Eur J Neurosci* **19**: 755.

Lijffijt M, Kenemans JL, Verbaten MN, van Engeland H (2005). A meta-analytic review of stopping performance in attention-deficit/hyperactivity disorder: deficient inhibitory motor control? *J Abnormal Psychol* **114**: 216–222.

Llorente AM, Voigt RG, Jensen CL, Berretta MC, Fraley JK, Heird WC (2006). Performance on a visual sustained attention and discrimination task is associated with urinary excretion of norepineprhine metabolite in children with attention-deficit/hyperactivity disorder (AD/HD). *Clin Neuropsychol* **20**: 133–144.

Luman M, Oosterlaan J, Sergeant JA (2005). The impact of reinforcement contingencies on AD/HD: a review and theoretical appraisal. *Clin Psychol Rev* **25**: 183–213.

Madras BK, Miller GM, Fischman AJ (2005). The dopamine transporter and attention-deficit/hyperactivity disorder. *Biol Psychiatry* **57**: 1397–1409.

Martinussen R, Hayden J, Hogg-Johnson S, Tannock R (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* **44**: 377–384.

Mazei MS, Pluto CP, Kirkbride B, Pehek EA (2002). Effects of catecholamine uptake blockers in the caudate-putamen and subregions of the medial prefrontal cortex of the rat. *Brain Res* **936**: 58–67 (DAT blockade preferentially increases BG rel to PFC DA).

McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004). Separate neural systems value immediate and delayed rewards. *Science* **306**: 503–507.

McClure S, Gilzenrat M, Cohen J (2006). An exploration-exploitation model based on norepinepherine and dopamine activity. *Adv Neural Inform Process Systems* **18**: 867–874.

Michelson D, Faries D, Wernicke J, Kelsey D, Kendrick K, Sallee FR et al (2001). Atomoxetine in the treatment of children and adolescents with attention-deficit/hyperactivity disorder: a randomized, placebo-controlled, dose–response study. *Pediatrics* **108**: E83.

Mobini S, Body S, Ho MY, Bradshaw CM, Szabadi E, Deakin JF et al (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology* **160**: 290–298.

Nigg JT, Casey BJ (2005). An integrative theory of attention-deficit/hyperactivity disorder based on the cognitive and affective neurosciences. *Dev Psychopathol* **17**: 785–806.

O'Reilly RC, Frank MJ (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput* **18**: 283–328.

Overtoom CC, Verbaten MN, Kenemans JL, van Engeland H, Buitelaar JK, van der Molen MW et al (2003). Effects of methylphenidate, desipramine, and L-dopa on attention and inhibition in children with attention deficit hyperactivity disorder. *Behav Brain Res* **145**: 7–15.

Owen AM, Doyon J, Dagher A, Sadikot A, Evans AC (1998). Abnormal basal ganglia outflow in Parkinson's disease identified with PET implications for higher cortical functions. *Brain* **121**: 949–965.

Redgrave P, Prescott TJ, Gurney K (1999). Is the short-latency dopamine response too short to signal reward error? *Trends Neurosci* **22**: 146–151.

Rhodes SM, Coghill DR, Matthews K (2005). Neuropsychological functioning in stimulant-naive boys with hyperkinetic disorder. *Psychol Med* **35**: 1109–1120.

Rinne UK, Laihinen A, Rinne JO, Nagren K, Bergman J, Ruotsalainen U (1990). Positron emission tomography demonstrates dopamine D2 receptor supersensitivity in the striatum of patients with early Parkinson's disease. *Mov Disord* **5**: 55–59.

Robbins TW, Everitt BJ (1999). Drug addiction: bad habits add up. *Nature* **398**: 567–570.

Ruskin DN, Bergstrom DA, Shenker A, Freeman LE, Baek D, Walters JR (2001). Drugs used in the treatment of attention-deficit/hyperactivity disorder affect postsynaptic firing rate and oscillation without preferential dopamine autoreceptor action. *Biol Psychiatry* **49**: 340–350.

Ruskin DN, Bergstrom DA, Tierney PL, Walters JR (2003). Correlated multisecond oscillations in firing rate in the basal ganglia: modulation by dopamine and the subthalamic nucleus. *Neuroscience* **117**: 427–438.

Russell V, Allie S, Wiggins T (2000). Increased noradrenergic activity in prefrontal cortex slices of an animal model for attention-deficit hyperactivity disorder—the spontaneously hypertensive rat. *Behav Brain Res* **117**: 69–74.

Russell VA, Wiggins TM (2000). Increased glutamate-stimulated norepinephrine release from prefrontal cortex slices of spontaneously hypertensive rats. *Metab Brain Dis* **15**: 297–304.

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

Sagvolden T, Johansen EB, Aase H, Russell VA (2005). A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes. *Behav Brain Sci* **28**: 397–419; discussion 419–68.

Santamaria A, O'Reilly R, Willcutt E (in preparation). Working memory deficits explain timing impairments in ADHD.

Scheres A, Dijkstra M, Ainslie E, Balkan J, Reynolds B, Sonuga-Barke E, Castellanos FX (2006a). Temporal probabilistic discounting of rewards in children adolescents: effects of age ADHD symptoms. *Neuropsychologia* **44**: 2092–2103.

Scheres A, Milham MP, Knutson B, Castellanos FX (2006b). Ventral striatal hyporesponsiveness during reward anticipation in attention deficit/hyperactivity disorder. *Biol Psychiatry* [Epub ahead of print].

Schiffer WK, Volkow ND, Fowler JS, Alexoff DL, Logan J, Dewey SL (2006). Therapeutic doses of amphetamine or methylphenidate differentially increase synaptic and extracellular dopamine. *Synapse* **59**: 243–251 (ritalin and amphetamine both increase synaptic (phasic) DA to same degree (both amp increases extracellular DA much more)).

Schultz W (2002). Getting formal with dopamine and reward. *Neuron* **36**: 241–263.

Seeman P, Madras B (2002). Methylphenidate elevates resting dopamine which lowers the impulse-triggered release of dopamine: a hypothesis. *Behav Brain Res* **130**: 79–83.

Servan-Schreiber D, Cohen JD, Steingard S (1997). Schizophrenic deficits in the processing of context: A test of a theoretical model. *Arch Gen Psychiatry* **53**: 1105–1113.

Sesack SR, Hawrylak VA, Matus C, Guido MA, Levey AI (1998). Dopamine axon varicosities in the prelimbic division of the rat

prefrontal cortex exhibit sparse immunoreactivity for the dopamine transporter. *J Neurosci* **18**: 2697–2708.

Solanto MV (2002). Dopamine dysfunction in AD/HD: integrating clinical and basic neuroscience research. *Behav Brain Res* **130**: 65–71.

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

Sonuga-Barke EJS (2003). The dual pathway model of AD/HD: an elaboration of neuro-developmental characteristics. *Neurosci Biobehav Rev* **27**: 593–604.

Sonuga-Barke EJS (2005). Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biol Psychiatry* **57**: 1231–1238.

Stryjer R, Klein C, Treves TA, Rabey JM (2005). The effects of acute loading with levodopa and levodopa with selegiline on blood pressure and plasma norepinephrine levels in chronic Parkinson's disease patients. *Acta Neurol Scand* **111**: 89–94.

Swanson CJ, Perry KW, Koch-Krueger S, Katner J, Svensson KA, Bymaster FP (2006). Effect of the attention deficit/hyperactivity disorder drug atomoxetine on extracellular concentrations of norepinephrine and dopamine in several brain regions of the rat. *Neuropharmacology* **50**: 755–760.

Todd RD, Huang H, Smalley SL, Nelson SF, Willcutt EG, Pennington BF *et al* (2005). Collaborative analysis of DRD4 and DAT genotypes in population-defined ADHD subtypes. *J Child Psychol Psychiatry* **46**: 1067–1073.

Toplak ME, Jain U, Tannock R (2005). Executive and motivational processes in adolescents with Attention-Deficit-Hyperactivity Disorder (ADHD). *Behav Brain Funct* **1**: 8.

Usher M, Cohen JD, Servan-Schreiber D, Rajkowski J, Aston-Jones G (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science* **283**: 549–554.

Volkow ND, Wang GJ, Fowler JS, Logan J, Gerasimov M, Maynard L *et al* (2001). Therapeutic doses of methylphenidate significantly increase extracellular dopamine in the human brain. *J Neurosci* **21**: RC121.

Weiner I, Joel D (2002). Dopamine in schizophrenia: Dysfunctional processing in basal ganglia -thalamocortical split circuits. In: Di Chiara G (ed). *Handbook of Experimental Pharmacology, Vol. 154/ii. Dopa1e in the CNS II*. Springer-Verlag: Berlin.

Wichmann T, Kliem MA, Soares J (2002). Slow oscillatory discharge in the primate basal ganglia. *J Neurophysiol* **87**: 1145–1148.

Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biol Psychiatry* **57**: 1336–1346.

Wu Q, Reith ME, Walker QD, Kuhn CM, Caroll FI, Garris PA (2002). Concurrent autoreceptor-mediated control of dopamine release and uptake during neurotransmission: an *in vivo* voltammetric study. *J Neurosci* **22**: 6272–6281.

Yano M, Steiner H (2005). Methylphenidate (Ritalin) induces Homer 1a and zif 268 expression in specific corticostriatal circuits. *Neuroscience* **132**: 855–865.

Yano M, Beverley JA, Steiner H (2006). Inhibition of methylphenidate-induced gene expression in the striatum by local blockade of D1 dopamine receptors: interhemispheric effects. *Neuroscience* **140**: 699–709.

Zametkin AJ, Rapoport JL (1987). Neurobiology of attention deficit disorder with hyperactivity: where have we come in 50 years? *J Am Acad Child Adolesc Psychiatry* **26**: 676–686.