

Neural and behavioral correlates of expectancy violations in attention-deficit hyperactivity disorder

Sarah Durston,^{1,2} Matthew C. Davidson,^{1,3} Martijn J. Mulder,² Julie A. Spicer,¹ Adriana Galvan,¹ Nim Tottenham,¹ Anouk Scheres,⁴ F. Xavier Castellanos,⁵ Herman van Engeland,² and B.J. Casey¹

¹Sackler Institute for Developmental Psychobiology, Weill Medical College of Cornell University, USA;

²Rudolf Magnus Institute of Neuroscience, Department of Child and Adolescent Psychiatry, University Medical Center Utrecht, The Netherlands; ³Department of Psychology, University of Massachusetts, USA; ⁴Department of Psychology, University of Arizona, USA; ⁵Child Study Center, New York University, USA

Background: Attention deficit hyperactivity disorder (ADHD) is a prevalent neuropsychiatric disorder in childhood with established problems in cognitive control and associated fronto-striatal circuitry. More recently, fronto-cerebellar circuits have been implicated in this disorder. Both of these circuits are important in predicting the occurrence and timing of behaviorally relevant events and in detecting violations of these predictions. Therefore, we hypothesized that the ability to predict the occurrence of frequent events would be compromised in ADHD, as well as the ability to adapt behavior when expectancy was violated. **Methods:** We used rapid, mixed-trial, event-related functional magnetic resonance imaging (fMRI) to examine cognitive and neural processes in two independent samples of children and adolescents with ADHD and matched controls. Subjects performed a variation of a go-no/go task where the predictability of stimulus identity (what) and timing (when) was manipulated. **Results:** Behaviorally, children and adolescents with ADHD had increased variability in reaction times, and decreased benefit in reaction time when events were predictable. Differences in accuracy between groups were most reliable for temporally unpredictable trials. Functional imaging results from both samples showed that relative to the control children and adolescents, individuals with ADHD had diminished cerebellar activity to violations of stimulus timing and diminished ventral prefrontal and anterior cingulate activity to violations in stimulus timing and identity. **Conclusions:** These findings are consistent with the view that disruptive behaviors in inappropriate contexts, a major criterion in diagnosing ADHD, may be related to an impaired ability to predict temporal and contextual cues in the environment, thus hindering the ability to alter behavior when they change. This ability requires intact fronto-cerebellar, as well as fronto-striatal circuitry. **Keywords:** ADD/ADHD, fMRI, expectancy, fronto-cerebellar.

Attention deficit hyperactivity disorder (ADHD) is characterized by developmentally inappropriate levels of activity and impulsivity, as well as disorganized and inattentive behaviors. This disorder has been associated with deficits in cognitive control, defined as the ability to suppress inappropriate behaviors over time and across different contexts (Nigg & Casey, 2005). Diagnostic criteria for ADHD reflect these deficits, as they include inappropriate and disruptive behaviors in inappropriate contexts, such as blurting out (unintentionally hurtful) comments, disruptive behavior in the classroom, and potentially dangerous behaviors, such as running out into traffic. Neuroimaging studies suggest that these deficits in cognitive control are associated with reduced activation in fronto-striatal circuitry (for review, Bush et al., 2005; Durston, 2003).

Cognitive control is dependent on the ability to predict temporal and contextual structure in the

environment (Nigg & Casey, 2005). Knowing *what* to expect and *when* to expect it is critical in planning and maintaining appropriate thoughts and actions over time. If basic expectations about the environment are inaccurate or less salient for children with ADHD, then their ability to detect *violations* of expectation will be compromised, and their ability to adjust their behavior accordingly will be weakened (Nigg & Casey, 2005). Specifically, if children with ADHD have trouble predicting the occurrence of events accurately, their reaction to a predictable event should be (1) slower and (2) more variable. Indirect support for this hypothesis comes from observations that children with ADHD have more variable response times and poor performance on tasks requiring time estimation (e.g., duration discrimination, finger tapping, and duration reproduction) to predictable events (Rubia et al., 2003; Toplak & Tannock, 2005; Toplak et al., 2006). In this study we tested whether children with ADHD are: (1) less able to benefit from trials being

Conflict of interest statement: No conflicts declared.

© 2007 The Authors

Journal compilation © 2007 Association for Child and Adolescent Mental Health.

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

predictable; and (2) less able to consequently adapt their behavior, both in terms of performance on the task and recruitment of relevant neural circuitry.

Previous studies have shown that learning to predict what and when events in the environment will occur depend on fronto-striatal (Schultz et al., 1997; Berns, Cohen, & Mintun, 1997; McClure et al., 2003; Zink et al., 2003; Rodriguez, Aron, & Poldrack, in press) and fronto-cerebellar (Keele & Ivry, 1990; Tesche & Karhu, 2000; Dreher & Grafman, 2002; Spencer et al., 2005) circuitry. Previously, we showed that activation in fronto-striatal regions was associated with violations of what to expect (Davidson et al., 2004), while activation in fronto-cerebellar areas was associated with violations in the timing of events (Casey, Amso, & Davidson, in press; Davidson et al., 2005).

MRI-based studies have shown the importance of fronto-striatal and fronto-cerebellar circuits in ADHD (e.g., Durston et al., 2004), as these regions show subtle decreases in volume (for review, Durston 2003; Seidman et al., 2005). A preliminary diffusion tensor imaging study showed decreased connectivity (fractional anisotropy) in pathways leading to and including the cerebellar peduncles (Ashtari et al., 2005). Traditionally, functional imaging studies of cognitive control in ADHD have shown changes in prefrontal and striatal activation (for review, Bush et al., 2005; Durston, 2003), while behavioral and electrophysiological studies deficits in temporal processing in ADHD have implicated cerebellum (e.g., Smith et al., 2002; Rubia et al., 2003; Van Meel et al., 2005). More recently, imaging studies have also shown changes in cerebellum (Zang et al., in press) and in response to stimulant medication (Anderson et al., 2002; Schweitzer et al., 2003), underscoring the potential relevance of this region to ADHD.

In this study, we hypothesized that the ability to predict the occurrence of events would be compromised in ADHD, as well as the ability to adapt behavior on trials where expectancy was violated. To test these hypotheses, we examined neural correlates underlying the predictability of stimulus identity (what) and timing (when) in two separate samples of children and adolescents with and without ADHD. Specifically, subjects participated in a variation of a go-no/go task where the predictability of events was manipulated in two ways: Expected or unexpected stimuli (go and no/go) were presented at expected or unexpected times. Based on previous work, we predicted regional differences in activation, with ventral prefrontal areas showing more sensitivity to stimulus violations (i.e., no/go-trials) and cerebellum showing more sensitivity to temporal violations in control subjects (Casey et al., in press; Davidson et al., 2004). Furthermore, we predicted a reduction in activation of these circuits during violations of stimulus or timing expectancy in subjects with ADHD compared to matched controls. We

collected data from two separate samples to provide an independent replication of the data. Samples differed in composition in terms of gender (80% vs. 100% male), age (11.9 vs. 14.9 years), IQ (98 vs. 108), the community from which they were recruited (New York City vs. Utrecht, the Netherlands) and MR-scanner used to acquire the data (GE 1.5T vs. Philips Allegra 1.5T), but both consisted of children and adolescents with ADHD and matched controls.

Methods

Participants

A total of 44 children and adolescents successfully completed the current study. Demographic information is provided in Table 1. Controls were matched to subjects with ADHD for age, gender, IQ, and hand preference in the combined sample, as well as in the two samples independently.

Sample 1 was acquired at the Sackler Institute for Developmental Psychobiology in New York and included ten children and adolescents with ADHD and ten matched controls. Seven additional children were scanned and excluded, either due to excessive head motion or failing to complete the experiment (4 subjects with ADHD, 3 controls). All subjects were screened for current or previous medical, neurological or psychiatric illnesses. For subjects with a clinical diagnosis of ADHD, a research diagnosis was established by K-SADS interview (Schedule for Affective Disorders and Schizophrenia for School-Age Children – Parent version; Orvaschel, 1994). Subjects that did not meet criteria for

Table 1 Descriptive variables per group (mean (standard deviation) and range)

	Controls	Subjects with ADHD
Sample 1		
GENDER (M/F)	8/2	8/2
AGE (years)	11.9 (2.1) 7.9–15.1	11.6 (2.6) 7.3–15.8
IQ	100 (10) 83–117	95 (12) 75–107
Hand preference (no. left-handed)	0/10	1/10
ADHD (no. meeting criteria)	0/10	5/10 combined 3/10 hyperactive 2/10 inattentive
ODD (no. meeting criteria)	0/10	4/10
Sample 2		
GENDER (M/F)	12/0	12/0
AGE (years)	15.0 (2.1) 11.9–19.7	14.9 (2.3) 9.6–17.6
IQ	107 (20) 76–150	108 (22) 84–150
Hand preference (no. left-handed)	3/12	3/12
ADHD (no. meeting criteria)	0/12	9/12 combined 3/12 hyperactive
ODD (no. meeting criteria)	0/12	4/12

ADHD, or that met criteria for any other disorder than ADHD or oppositional defiant disorder (ODD), were excluded. Control subjects were required to be free of any diagnosis on K-SADS interview. All subjects participated in IQ-testing using the WASI (Wechsler, 1999). None of the subjects were diagnosed with learning disorders.

Sample 2 was recruited as part of an ongoing, larger study at the University Medical Center Utrecht in the Netherlands and consisted of twelve boys with ADHD and twelve matched controls. Subjects with major physical or neurological illness were excluded. All subjects participated in an IQ assessment using the Wechsler Intelligence Scale for Children – Revised (WISC-R; Wechsler, 1974) and a parent participated in a semi-structured interview session to confirm or disconfirm the clinical psychiatric diagnosis using the Diagnostic Interview Schedule for Children (DISC-P; Shaffer et al., 2000). ADHD subjects were required to meet DSM-IV criteria for ADHD, as assessed by DISC-interview. Subjects with co-morbid disorders other than ODD were excluded. Control subjects were excluded if they met DSM-IV criteria for any psychiatric diagnosis, as assessed by DISC interview. None of the subjects were diagnosed with learning disorders.

At both sites, subjects with ADHD on stimulant medication discontinued treatment for a minimum of 24 hours prior to the scan. Only subjects on short-acting stimulants were included. Fourteen subjects with ADHD were on stimulant medication (5/10 in *sample 1* and 9/12 in *sample 2*). Participants were acclimated to the scanning environment with a simulator-system prior to the actual scanning session. After complete description of the study, informed assent/consent was obtained from all subjects and their parents. All procedures were conducted in accordance with the guidelines established by the Institutional Review Board at the relevant institution.

Behavioral paradigm

The task used at each site was identical in structure but varied in number of trial types.¹ Subjects were asked to perform a visual target detection task by pressing a single response button with their right index finger whenever a target-stimulus was presented (Davidson et al., 2004). The task was a variation of a go/no/go task, where the temporal predictability of events was manipulated in addition to the predictability of the identity of the stimulus (i.e., no/go vs. go-stimuli). It was designed to build up the expectancy of an event occurring (the frequent and predictable go-trial). Subjects were required to adjust their behavior when that prediction was violated (i.e., inhibit a prepotent button-press on no/go-trials or press the button at an unexpected time on unpredictable go-trials). This design allowed us to assess whether subjects were building up an expectancy of a go-trial at a predictable time by

assessing whether they were faster on predictable than on unpredictable go-trials. Furthermore, we were able to investigate their ability to consequently alter their behavior on unpredictable trial types.

The task was presented in the context of a computer game, where subjects were asked to 'help feed a hungry little mouse as much cheese as possible'. The target-stimulus was a cartoon drawing of a piece of cheese (go-trial), whereas the unexpected stimulus was a cartoon drawing of a cat (no/go-trial). During the interstimulus interval, a mouse-hole remained on the screen, briefly opening to reveal one of the experimental stimuli in a continuous stream of trials (Davidson et al., 2004).

The task involved four blocks of 72 trials (288 trials total). On the majority of trials (76% of 288 = 192 trials), the target-stimulus (cheese) was presented at the expected time (every 4 s). On the remainder of the trials (24% of 288 = 96 trials), unpredictable trials (24 of each type) were presented: (1) an unexpected stimulus at the expected time (cat at 4 s; temporally predictable no/go-trial); (2) an expected target stimulus at an unexpected time (cheese at 2 s; temporally unpredictable go-trial); and (3) an unexpected stimulus at the unexpected time (cat at 2 s; temporally unpredictable no/go-trial). A fourth unpredictable trial type, where no stimulus was presented at the predictable time (a non-trial), was also included to help prevent subjects learning any pattern other than that of the predictable go-trials. Trial types were mixed pseudo-randomly in equal numbers throughout each block. The stimuli were presented for 500 ms, with an ISI of either 1500 or 3500 ms.

Analysis of behavioral data

All behavioral data were analyzed using the SPSS statistical package (version 14.0). Accuracy scores were calculated as the percentage of hits to targets (predictable and unpredictable go-trials) and correct omissions to non-targets (no/go-trials) relative to the total number of trials of each type. Mean reaction times (RT) were calculated for the condition for which a response was required (i.e., expected stimulus at the expected or unexpected time). Differences in behavior were investigated for the combined sample, as well as separately for each independent sample. Subjects' ability to build up expectancy was investigated using two measures: (1) benefit in reaction time on temporally predictable compared to temporally unpredictable trials (mean RT on unpredictable go-trials minus mean RT on predictable go-trials) and (2) variability in RT on go-trials.

Image acquisition

In New York, EPI-BOLD images were acquired in a 1.5T GE scanner (General Electric Medical Systems, Milwaukee, Wisconsin) with a quadrature head coil, using a protocol of TR = 2000, TE = 40, and flip angle = 90 degrees. Twenty-four coronal slices of 6 mm slice thickness and zero gap (3.125 × 3.125 m in-plane resolution) were collected for 144 repetitions (including 4 discarded acquisitions). Anatomical images (TR = 500, TE = minimum, 6 mm slice thickness) were collected in the same locations as the functional images, as

¹ In *sample 2*, two subjects with ADHD and three control subjects performed a slightly different version of the task, without trials at unexpected times. Therefore, these analyses from this site include 10 subjects with ADHD and 9 controls. For the other subjects in *sample 2*, the paradigm was identical to the one used for *sample 1*.

well as a set of high-resolution SPGR images (three-dimensional spoiled gradient, 256×256 in-plane resolution, 240 mm field of view, 124×1.5 mm coronal slices with TR = 25 and TE = 5) that were used for 3D localization.

In Utrecht, MRI images were acquired on a 1.5-T Philips Allegra (Philips Medical Systems, Best, the Netherlands). Functional MRI scans consisted of a navigated three-dimensional (3D)-PRESTO pulse sequence (TE 11 ms, TR 21.74 ms, flip angle 9.0° , matrix $64 \times 64 \times 36$, FOV $256 \times 256 \times 144$ mm, voxel size 4 mm isotropic, and scan duration 2.0 s per 36-slice volume), covering the whole brain. Anatomical T1-weighted 3D fast field echo (FFE) scans with 170 1.2 mm contiguous coronal slices of the whole head (TE 4.6 ms, TR 30 ms, flip angle 30° , FOV 256 mm, in plane voxel size 1 mm \times 1 mm) were also acquired. An FA30 scan with contrast more similar to the T1 weighted scans was also acquired to aid in the alignment of PRESTO images to the template (TE 11 ms, TR 21.74 ms, flip angle 30° , matrix $64 \times 64 \times 36$, FOV $256 \times 256 \times 144$ mm, voxel size 4 mm isotropic).

Image analysis

Data were analyzed separately for both sites, using a random effects model in Statistical Parametric Mapping software (SPM2; Wellcome Department of Imaging Neuroscience, London, UK). The fMRI data were not pooled across sites in order to: (1) allow us to independently replicate our findings within the context of this study; and (2) avoid issues associated with pooling data from different MR-scanners and pulse sequences. For both analyses, functional MR-images were re-aligned and normalized to a standard stereotactic space (Montreal Neurological Institute (MNI) template) and spatially smoothed with a Gaussian kernel size of two voxels (6 and 8 mm at FWHM for *sample 1* and *sample 2*, respectively). Estimated motion parameters were examined on a subject-by-subject basis to ensure that the amount of motion did not exceed the size of 1 voxel.

At the first level, eleven event types were defined (initial fixation with no behavioral output, expected and unexpected stimuli at expected and unexpected times, and the omission of an event, all of which could be correct or incorrect). These included four effects of interest (correct trials for expected and unexpected stimuli at the expected and unexpected time). The event types were time-locked to the stimulus by a canonical synthetic haemodynamic response function (HRF) and its first-order temporal derivative.

At the second level, three different analyses were performed. First, a one-sample *t*-test was performed separately for both samples of control subjects to compare our findings from both sites. Only correct trials were included in the analyses. Here, three conditions of interest were investigated, comparing unpredictable events that required subjects to adapt their behavior to predictable events (expected stimulus at unexpected time (temporally unpredictable go-trial); unexpected stimulus at expected time (temporally predictable no/go-trial); unexpected stimulus at unexpected time (temporally unpredictable no/go-trial)). A threshold of $p < .005$, with a minimum extent of 10

voxels, was used in each case to minimize the rate of false positive clusters (Forman et al., 1995).

Second, an ROI-analysis was performed to investigate differences between controls and subjects with ADHD for the combined sample. Here, data were pooled from both sites, as ROI-analyses have the advantage of clustering signal from numerous voxels, thereby controlling for the local effects that could be problematic in a voxel-level analysis (e.g., differences in voxel size, direction of scan acquisition). For example, differences in scan direction could differentially affect susceptibility artifacts in critical regions such as IFG. ROIs were chosen based on *a-priori* hypotheses of differences in ventral prefrontal and cerebellar regions: Three ROIs were investigated, right inferior frontal gyrus (rIFG), anterior cingulate gyrus (ACG) and cerebellum (CB). Regions were functionally defined using the MarsBaR package (Brett et al., 2002) and were taken from one-sample *t*-tests for control subjects.

Finally, an exploratory voxel-by-voxel investigation of differences between subjects with ADHD and controls was conducted. Here, whole-brain, two-sample *t*-tests were performed for each sample separately, investigating three conditions of interest (expected stimulus at unexpected time; unexpected stimulus at expected time; unexpected stimulus at unexpected time). A threshold of $p < .005$, with a minimum extent of 5 voxels, was used (Forman et al., 1995). Again, only correct trials were included in the analyses.

Results

Behavioral results

Combined sample. In the combined analysis of 22 subjects with ADHD and 22 controls, there were no significant differences in overall RT, although there were significant differences in RT variability and RT benefit on temporally predictable trials (see Figure 1). Subjects with ADHD were more variable and benefited less from trials being temporally predictable (standard deviation in RT = 93 (31) ms for controls; 131 (38) ms for ADHD; $t(42) = 3.75$; $p = .001$; RT benefit = 76 (27) ms for controls; 57 (33) ms for ADHD; $t(42) = 1.95$; $p = .05$).

Accuracy was significantly lower for subjects with ADHD compared to controls for manipulations of timing (expected and unexpected stimuli at unexpected time), as well as manipulations of stimulus type (unexpected stimuli at expected and unexpected time) (all $t(42) > 2.1$; $p < .04$). Both omission and commission errors contributed to the timing accuracy score and both differed significantly between groups ($t > 2.4$; $p > .02$).

Individual samples. Reaction time results were similar for the two independent samples, as there were no differences in overall RT, but there were significant differences in RT variability (*Sample 1*: standard deviation in RT = 106 (36) ms for controls; 149 (33) ms for ADHD; $t(18) = 2.76$; $p = .01$; *sample 2*: standard deviation in RT = 81 (21) ms for controls;

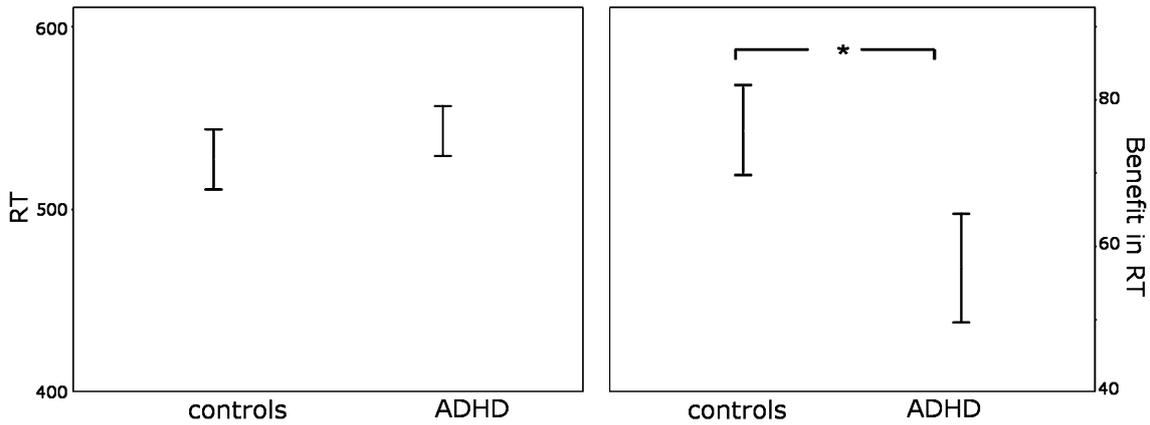


Figure 1 Overall reaction time was not different between groups, but benefit in reaction time when trials were temporally predictable was. *significant difference between both groups ($p < .05$)

117 (35) ms for ADHD; $t(22) = 2.99$; $p = .007$). There was also a significant difference in RT benefit on predictable go-trials in *sample 1* (controls = .79 (27) ms; ADHD = 47 (31) ms; $t = 2.50$; $p = .02$).

The most reliable difference in accuracy between diagnostic groups was for the timing manipulation, as differences here were significant in both samples (*Sample 1*: controls = 88 (9)% vs. ADHD = 78 (10)%; *sample 2*: controls = 92 (8)% vs. ADHD = 78 (11)%; $t(42) > 2.4$; $p < .03$). Figure 2 shows accuracy for timing manipulations from both sites. For *sample 2*, differences in accuracy also reached significance for manipulations of stimulus identity (controls = 77 (20)% vs. ADHD = 56 (19)%; $t(22) = 2.3$; $p = .03$).

Functional MRI results

Results for control subjects are summarized in Table 2 for both samples. Findings were similar to

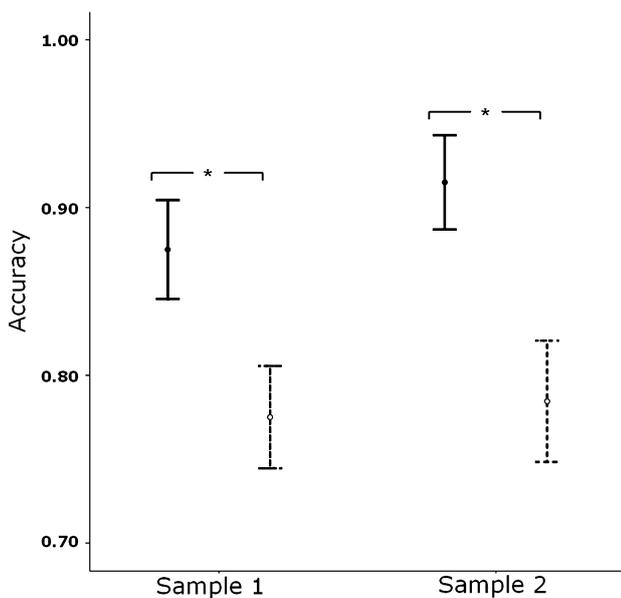


Figure 2 The most reliable differences between diagnostic groups were on temporally unpredictable trials in both samples. *significant difference between groups ($p < .03$)

previous results with this task (Davidson et al., 2005) in that cerebellum was particularly sensitive to violations of timing (expected stimulus at the unexpected time) (see Figure 3), and regions in ventral prefrontal cortex and anterior cingulate cortex were particularly sensitive to the manipulation of stimulus type (unexpected stimulus at expected and unexpected times) (see Figure 4). Cerebellar results only reached significance for one side in each sample, but in both samples, subthreshold bilateral activation was present.

An ROI-analysis was conducted for the combined sample, for three, functionally defined regions of *a-priori* interest, anterior cingulate gyrus (ACG), right inferior frontal gyrus (rIFG) and cerebellum (CB). For the manipulation of stimulus type (unexpected stimulus at expected time; no/go-trial), activation was increased for control subjects compared to subjects with ADHD in ACG ($t = 3.02$; $df = 42$; $p = .004$; Figure 5). For the manipulation of stimulus timing (expected stimulus at unexpected time), activation was increased for control subjects compared to subjects with ADHD in CB ($t = 3.97$; $df = 37$; $p = .0003$; Figure 5). For the combined manipulation of stimulus type and timing (unexpected stimulus at unexpected time), differences between groups did not reach significance for any of the three ROIs.

Exploratory, whole-brain, two-sample *t*-tests showed greater activation for control subjects than subjects with ADHD in a limited number of regions (see Table 2): For *sample 1*, differences in hypothesized regions only reached significance for the combined manipulation of stimulus type and timing (unexpected stimulus at unexpected time). Here, significantly greater activation was found for control children than children with ADHD in right inferior frontal gyrus (BA 45; talairach: 53, 3, 5; $t = 3.82$; $df = 18$; $p < .005$; min extent = 5 voxels). For *sample 2*, differences reached significance for the manipulations of timing (expected stimulus at unexpected time) and stimulus type (unexpected stimulus at expected time). For the manipulation of timing,

Table 2 fMRI results for manipulations of stimulus type and timing. Regions are taken from the one-sample *t*-test for control subjects in *sample 1* ($t(18) > 3.25$; $p < .005$; $k > 10$ voxels) and *sample 2* ($t(22) > 3.11$; $p < .005$; $k > 10$ voxels)

Region	Side	BA	Talairach	Max <i>t</i> -value
Expected stimulus, unexpected time (temporally unpredictable go-trial)				
<i>Sample 1</i>				
cerebellum *	R		22, -58, -7	7.25
<i>Sample 2</i>				
cerebellum † *	L		-20, -52, -44	5.59
IFG	R	45	40, 20, -4	4.40
MTG	R	21	44, -23, -2	7.83
LG	med	18	0, -82, 4	9.42
Unexpected stimulus, expected time (temporally predictable no/go-trial)				
<i>Sample 1</i>				
STG	R	21	62, -27, 1	4.41
<i>Sample 2</i>				
IFG †	R	45/47	48, 23, -8	5.62
	L	45/47	-48, 20, -1	4.37
ACG † *	med	32	8, 41, 50	7.57
MFG	R	46	40, 48, 23	6.02
		6	44, 10, 51	4.78
	L	46	-35, 29, 30	5.78
ITG	L	37	-42, -80, 1	5.01
MTG	R	21	44, -43, -8	7.36
STG	R		59, -19, 1	5.64
Unexpected stimulus, unexpected time (temporally unpredictable no/go-trial)				
<i>Sample 1</i>				
IFG †	R	45	59, 22, 10	4.62
ACG	med	32	0, 34, 26	4.56
SFG	med	8	2, 45, 42	5.87
<i>Sample 2</i>				
IFG	R	47	40, 27, -5	5.98
ACG	med	32	4, 25, 42	9.07
SFG	R	8	36, 18, 51	6.53
IPL	L	40	-55, -50, 28	8.14

ACG = anterior cingulate gyrus; BA = Brodmann Area; G = gyrus; IFG = inferior frontal gyrus; IPL = inferior parietal lobule; ITG = inferior temporal gyrus; L = left; med = medial; MFG = middle frontal gyrus; MTG = middle temporal gyrus; R = right; SFG = superior frontal gyrus; STG = superior temporal gyrus; † significant difference in two-sample *t*-test; * significantly more activation for control subjects than subjects with ADHD in post-hoc ROI omnibus analysis; **BOLD font** indicates regions that were hypothesized for a given comparison.

In *sample 2*, two subjects with ADHD and three control subjects performed a slightly different version of the task, without trials at unexpected times. Therefore, these analyses from this site include 10 subjects with ADHD and 9 controls. For the other subjects in *sample 2*, the paradigm was identical to the one used for *sample 1*.

children with ADHD activated inferior cerebellum significantly less than control children (talairach: -16, -56, -41; $t = 4.50$; $df = 17$; $p < .005$; min extent = 5 voxels), whereas for the manipulation of stimulus type (temporally predictable no/go-trial), they activated anterior cingulate gyrus (BA 32; talairach: 20, 21, 32; $t = 3.49$; $df = 22$; $p < .005$; min extent = 5 voxels) and inferior frontal gyrus (BA 45; talairach: -36, 44, 9; $t = 4.01$; $df = 22$; $p < .001$; min extent = 5 voxels) significantly less.

Discussion

In this study, we examined how well individuals with ADHD could anticipate predictable events and consequently modify behavior when these predictions were violated. Reaction times on predictable trials were more variable for individuals with ADHD and they benefited less in reaction time from trials being predictable. Differences in accuracy between diagnostic groups were most reliable for trials that were temporally unpredictable and these differences were paralleled by reductions in cerebellar activity in individuals with ADHD across samples. Furthermore, individuals with ADHD also showed attenuated activation in prefrontal regions in response to violations of stimulus type. Taken together, these results suggest that children and adolescents with ADHD may not be predicting the occurrence of future events to the same degree as controls and therefore not able to prepare their response as well, leading to poorer performance on both predictable and unpredictable trials and changes in activity in fronto-cerebellar circuits.

In both samples, the control group of typically developing children and adolescents showed a pattern of brain activation in response to violations of expectancy that was similar to previous results with this task (see Table 2; Davidson et al., 2005). The replication of previous findings in two independent samples of subjects adds confidence in these results. In both samples, controls activated regions in ventral prefrontal cortex and anterior cingulate gyrus in response to violations of stimulus identity (at both predictable and unpredictable times), and regions in cerebellum to violations of stimulus timing (Figures 2 & 3).

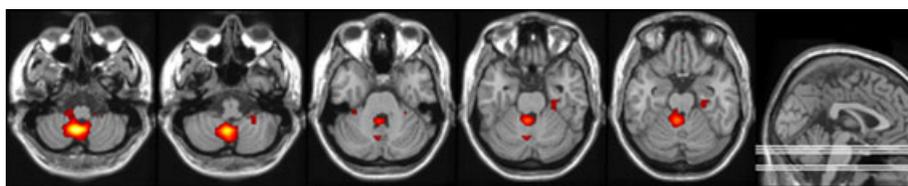


Figure 3 Typical pattern of cerebellar activity to violations of stimulus timing for controls where differences between diagnostic groups were shown. Figure is from *sample 2*; $t_9 > 3.11$; $p < .005$; minimum extent = 10 voxels.

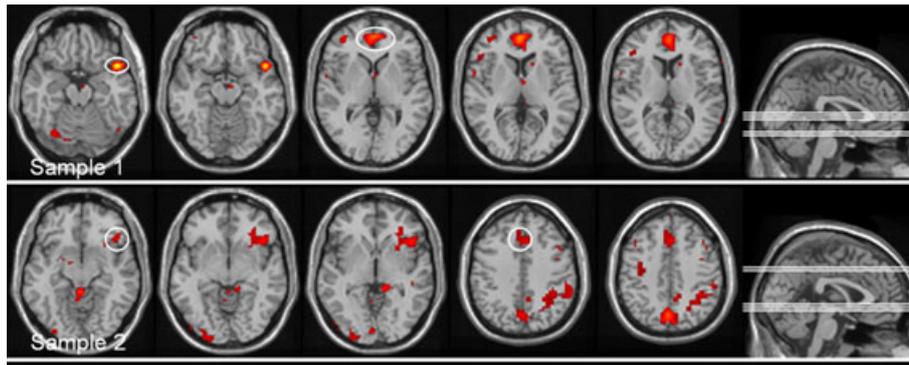


Figure 4 Typical pattern of RIFG and ACG activity replicated in both samples of control subjects. Maps represent activation patterns to stimuli where expectancy was violated in terms of stimulus identity and stimulus timing. $t_9 > 3.25$ for sample 1; $t_{11} > 3.11$ for sample 2; $p < .005$; $df = 9$; minimum extent 10 voxels.

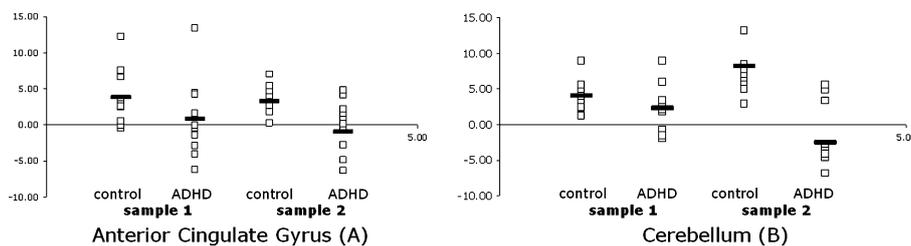


Figure 5 Activation in ACG during violations of stimulus identity (A) and in cerebellum for violations of stimulus timing (B), for both samples separately

Behaviorally, the most reliable difference between subjects with ADHD and controls was for violations of stimulus timing. This was paralleled by attenuated cerebellar activity for subjects with ADHD relative to controls (Figures 2 & 5). Typically, cognitive control paradigms require subjects with ADHD to suppress prepotent information or responses. In this paradigm, subjects with ADHD were not only required to suppress a behavioral response on trials where an unexpected stimulus occurred, but were also required to adapt their behavior on trials where an expected stimulus occurred at an unexpected time. The lower accuracy and attenuated cerebellar activation on these trials, together with reduced RT benefit on predictable trials suggest that cerebellar input to prefrontal cortical areas, involved in the prediction of events, may be weakened or impaired in this disorder. The ability to predict the occurrence of events is essential to planning and maintaining appropriate actions in different contexts, and if this ability is compromised in ADHD, poor behavioral adjustment in contexts where expectancy is violated may be in part related to this deficit.

In addition to changes associated with timing violations, children and adolescents with ADHD activated ventral prefrontal cortex and anterior cingulate gyrus less in response to violations of stimulus type. These results parallel previous reports of reductions in activation in prefrontal and anterior cingulate cortex in ADHD, associated with suppress-

ing a behavioral response or competing information (for review, Bush et al., 2005; Durston, 2003).

In addition to the similarities between our findings in both samples, we also report differences. For example, control subjects in *sample 1* had lower accuracy on trials where a button press needed to be suppressed (unexpected stimuli), as well as less activation in prefrontal areas on successful trials of this type. These differences may reflect differences in sample composition, as *sample 1* on average was younger than *sample 2* and had lower IQ. The pattern of lower prefrontal activation in younger subjects is in line with the literature on development of cognitive control that describes an increase of prefrontal involvement in this type of task with development (for review, see Durston & Casey, 2006). There were similar differences in activation between our two samples of ADHD-subjects that likely reflect differences in sample composition.

Other noticeable differences from Table 2 are a greater extent of activation in *sample 2* than *sample 1* and a limited overlap in regions activated. This may reflect a number of factors, in addition to developmental differences, including differences in sample size, and MR-scanners and sequences used.

There are limitations in our study that need to be acknowledged. First, the use of an adult template in a pediatric population could be problematic due to differences in gross brain anatomy. However, there is evidence to suggest that anatomic differences

between school-aged children and adults are only modest in their effects on detecting functional differences (Burgund et al., 2002). Furthermore, the age range of the samples included in this study spans the ages of 7 to 19 years, meaning that a child or adolescent template brain would have been equally problematic. Second, although subjects with ADHD were either not on medication, or discontinued treatment 24 hours prior to participating in the fMRI scan, most were not stimulant-naïve. As such, we cannot rule out that some of the observed changes in brain activation are due to long-term effects of stimulant medication.

Conclusions

We show that children and adolescents with ADHD have lower accuracy and reduced cerebellar activation for trials where stimulus timing is violated, in addition to reduced prefrontal activation to violations of stimulus type. This pattern is paralleled by a decrease in RT benefit when events are predictable. Knowing what to expect, and when to expect it are important cognitive processes and adjusting behavior when predictions are violated is an essential element of cognitive control, as this ability is critical for planning and maintaining appropriate thoughts and actions over time. Our findings are consistent with the view that disruptive behaviors in inappropriate contexts, a major criterion in diagnosing ADHD, may be related to an impaired ability to predict temporal and contextual cues in the environment, and alter behavior when they change. This ability requires intact fronto-cerebellar circuitry, as well as frontostriatal circuitry.

Acknowledgements

The authors gratefully acknowledge all the subjects and their families that participated in this study. This work was supported in part by a Dutch Science Foundation (NWO) VENI grant to SD, and by R01 MH63255 and R21 DA1588 to BJC.

Correspondence to

Sarah Durston, Rudolf Magnus Institute of Neuroscience, Department of Child and Adolescent Psychiatry, Neuroimaging Lab, HP A 01.468, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands; Tel: + 31 30 250 8161; Fax: + 31 30 250 5444; Email: S.Durston@umcutrecht.nl

References

Anderson, C.M., Polcari, A., Lowen, S.B., Renshaw, P.F., & Teicher, M.H. (2002). Effects of methylpheni-

- date on functional magnetic resonance relaxometry of the cerebellar vermis in boys with ADHD. *American Journal of Psychiatry*, *159*, 1322–1328.
- Ashtari, M., Kumra, S., Bhaskar, S.L., Clarke, T., Thaden, E., Cervellione, K.L., Rhinewine, J., Kane, J.M., Adelman, A., Milanaik, R., Maytal, J., Diamond, A., Szeszko, P., & Ardekani, B.A. (2005). ADHD: A preliminary diffusion tensor imaging study. *Biological Psychiatry*, *57*, 448–455.
- Berns, G.S., Cohen, J.D., & Mintun, M.A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science*, *276*, 1272–1275.
- Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). Region of interest analysis using an SPM toolbox [abstract]. *NeuroImage*, *16*.
- Burgund, E.D., Kang, H.C., Kelly, J.E., Buckner, R.L., Snyder, A.Z., Petersen, S.E., & Schlaggar, B.L. (2002). The feasibility of a common stereotactic space for children and adults in fMRI studies of development. *NeuroImage*, *17*, 184–200.
- Bush, G., Valera, E.M., & Seidman, L.J. (2005). Functional neuroimaging of attention-deficit/hyperactivity disorder: A review and suggested future directions. *Biological Psychiatry*, *57*, 1273–1284.
- Casey, B.J., Amso, D., & Davidson, M.C. (in press). Learning about learning and development with neuroimaging. *Attention and Performance XXI: Processes of Change in Brain and Cognitive Development*.
- Davidson, M.C., Horvitz, J.C., Tottenham, N., Fossella, J.A., Watts, R., Ulug, A.M., & Casey, B.J. (2004). Differential cingulate and caudate activation following unexpected nonrewarding stimuli. *NeuroImage*, *23*, 1039–1045.
- Davidson, M., Tottenham, N., Spicer, J., Galvan, A., Durston, S., Horvitz, J., & Casey, B.J. (2005). Neural responses to stimulus and temporal violations of expectation: A developmental fMRI study. *Journal of Cognitive Neuroscience*, (Suppl.) *S*, 191.
- Dreher, J.C., & Grafman, J. (2002). The roles of the cerebellum and basal ganglia in timing and error prediction. *European Journal of Neuroscience*, *16*, 1609–1619.
- Durston, S. (2003). A review of the biological bases of ADHD: What have we learned from imaging studies? *Mental Retardation and Developmental Disabilities Research Reviews*, *9*, 184–195.
- Durston, S., & Casey, B.J. (2006). What have we learned about cognitive development from neuroimaging? *Neuropsychologia*, *44*, 2149–2157.
- Durston, S., Hulshoff, P.H.E., Schnack, H.G., Buitelaar, J.K., Steenhuis, M.P., Minderaa, R.B., Kahn, R.S., van Engeland, H. (2004). Magnetic resonance imaging of boys with ADHD and their unaffected siblings. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*, 332–340.
- Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., & Noll, D.C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI). *Magnetic Resonance Medicine*, *33*, 636–647.
- Keele, S.W., & Ivry, R. (1990). Does the cerebellum provide a common computation for diverse tasks? A timing hypothesis. *Annals of the New York Academy of Sciences*, *608*, 179–207.

- McClure, S.M., Daw, N.D., & Montague, P.R. (2003). A computational substrate for incentive salience. *Trends in Neurosciences*, *26*, 423–428.
- Nigg, J.T., & Casey, B.J. (2005). An integrative theory of attention-deficit/hyperactivity disorder based on the cognitive and affective neurosciences. *Developmental Psychopathology*, *17*, 785–806.
- Orvaschel, H. (1994). *Schedule for Affective Disorders and Schizophrenia for School Age Children – Parent version (KSADS-P) (5th edn)*. Fort Lauderdale, FL: Nova Southeastern University, Center for Psychological studies.
- Rodriguez, P.F., Aron, A.R., & Poldrack, R.A. (in press). Ventral-striatal/nucleus-accumbens sensitivity to prediction errors during classification learning. *Human Brain Mapping*.
- Rubia, K., Noorloos, J., Smith, A., Gunning, B., & Sergeant, J. (2003). Motor timing deficits in community and clinical boys with hyperactive behavior: The effect of methylphenidate on motor timing. *Journal of Abnormal Child Psychology*, *31*, 301–313.
- Seidman, L.J., Valera, E.M., & Makris, N. (2005). Structural brain imaging of ADHD. *Biological Psychiatry*, *57*, 1263–1272.
- Schultz, W., Dayan, P., & Montague, P.R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599.
- Schweitzer, J.B., Lee, D.O., Hanford, R.B., Tagamets, M.A., Hoffman, J.M., Grafton, S.T., & Kilts, C.D. (2003). A positron emission tomography study of methylphenidate in adults with ADHD: Alterations in resting blood flow and predicting treatment response. *Neuropsychopharmacology*, *28*, 967–973.
- Shaffer, D., Fisher, P., Lucas, C., Dulcan, M., & Schwab-Stone, M. (2000). NIMH Diagnostic Interview Schedule for Children version IV (NIMH DISC-IV). Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*, 28–38.
- Smith, A., Taylor, E., Rogers, J.W., Newman, S., & Rubia, K. (2002). Evidence for a pure time perception deficit in children with ADHD. *Journal of Child Psychology and Psychiatry*, *43*, 529–542.
- Spencer, R.M., Ivry, R.B., & Zelaznik, H.N. (2005). Role of the cerebellum in movements: Control of timing or movement transitions? *Experimental Brain Research*, *161*, 383–396.
- Tesche, C.D., & Karhu, J.J. (2000). Anticipatory cerebellar responses during somatosensory omission in man. *Human Brain Mapping*, *9*, 119–142.
- Toplak, M.E., Dockstader, C., & Tannock, R. (2006). Temporal information processing in ADHD: Findings to date and new methods. *Journal of Neuroscience Methods*, *151*, 15–29.
- Toplak, M.E., & Tannock, R. (2005). Time perception: Modality and duration effects in attention-deficit-hyperactivity disorder (ADHD). *Journal of Abnormal Child Psychology*, *33*, 639–654.
- Van Meel, C.S., Oosterlaan, J., Heslenfeld, D.J., & Sergeant, J.A. (2005). Motivational effects on motor timing in ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*, 451–460.
- Wechsler, D. (1974). *Wechsler Intelligence Scale for Children – Revised (WISC-R)*. New York: Psychological Corporation.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio TX: Psychological Corporation.
- Zang, Y.F., Jin, Z., Weng, X.C., Zhang, L., Zeng, Y.W., & Yang, L. (in press). Functional MRI in ADHD: Evidence for hypofrontality. *Brain Development*.
- Zink, C.F., Pagnoni, G., Martin, M.E., Dhamala, M., & Berns, G.S. (2003). Human striatal response to salient nonrewarding stimuli. *Journal of Neuroscience*, *23*, 8092–8097.

Manuscript accepted 8 February 2007