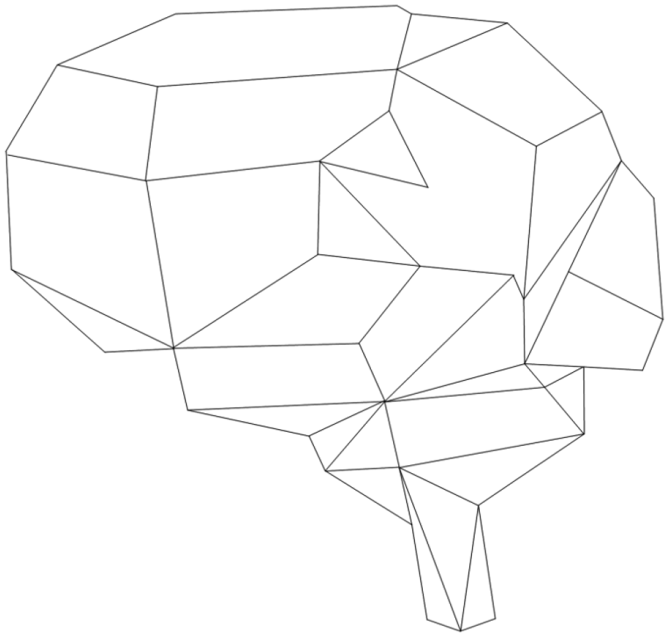
Stimulants and the developing brain

Schweren, Lizanne Johanna Stephanie

Chapter 5

STIMULANT TREATMENT TRAJECTORIES ARE ASSOCIATED WITH NEURAL REWARD PROCESSING IN ADHD



In press as:
Schweren LJS, Groenman A, von Rhein D, Weeda W, Faraone SV, Luman M, van Ewijk H, Heslenfeld DJ, Franke B, Buitelaar JK, Oosterlaan J, Hoekstra PJ, Hartman CA. Stimulant treatment trajectories are associated with neural reward processing in attention-deficit/hyperactivity disorder. J Clin Psychiatry, 2016.

ABSTRACT

Objective: The past decades have seen a surge in stimulant prescriptions for the treatment of attention-deficit/hyperactivity disorder (ADHD). Stimulants acutely alleviate symptoms and cognitive deficits associated with ADHD by modulating striatal dopamine neurotransmission, and induce therapeutic changes in brain activation patterns. Long-term functional changes after treatment are unknown, as long-term studies are scarce and have focused on brain structure. In this observational study (2009-2012), we investigated associations between lifetime stimulant treatment history and neural activity during reward processing.
Methods: Participants fulfilling DSM-5 criteria for ADHD (n=269) were classified according to stimulant treatment trajectory. Of those, 124 performed a monetary incentive delay task during magnetic resonance imaging, all in their non-medicated state ($N_{EARLY\&INTENSE}$=51; $N_{LATE\&MODERATE}$=49; $N_{EARLY\&MODERATE}$=9; N_{NAIVE}=15; mean age=17.4 years, range 10-26 years). Whole-brain analyses were performed with additional focus on the striatum, concentrating on the two largest treatment groups.
Results: Compared to the 'late-and-moderate' treatment group, the 'early-and-intense' treatment group showed more activation in the supplementary motor area and dorsal anterior cingulate cortex (SMA/dACC) during reward outcome (cluster size=8696 mm^3; $p_{CLUSTER}<0.001$). SMA/dACC activation of the control group fell in between the two treatment groups. Treatment history was not associated with striatal activation during reward processing.
Conclusions: Our findings are compatible with previous reports of acute increases of SMA/dACC activity in individuals with ADHD after stimulant administration. Higher SMA/dACC activity may indicate that patients with a history of intensive stimulant treatment, but currently off-medication, recruit brain regions for cognitive control and/or decision-making upon being rewarded. No striatal or structural changes were found.

CLINICAL POINTS

- Stimulant treatment is regarded a safe and effective treatment for ADHD symptoms, yet their long-term effects on brain activation patterns in children and adolescents are largely unknown.
- Early and intense stimulant treatment may result in increased activation of cognitive control areas during rewarding situations, even when patients are non-medicated at that time.

INTRODUCTION

Stimulant treatment is the medical intervention of first choice for children and adolescents with attention-deficit/hyperactivity disorder (ADHD). The past decades have seen a surge in stimulant prescription rates (Trip et al., 2009). Alleviation of symptoms and cognitive deficits associated with ADHD appears – in general – not to last after medication is discontinued, and there is little evidence of long-term improved functioning (Jensen et al., 2007; Molina et al., 2009; Van De Loo-Neus et al., 2011). The absence of conclusive evidence regarding potential long-term effects of stimulant treatment, either positive or negative, has unsettled parents, patients, and society at large.

Studies of long-term stimulant treatment effects on brain structure have yielded mixed results. Two meta-analyses found that striatal volume was more reduced in patients compared to controls when the ADHD sample included more treatment-naive patients (Frodl & Skokauskas, 2012; Nakao et al., 2011), suggesting that striatal volume reduction observed in ADHD is driven by untreated rather than stimulant-treated patients. However, a large-scale longitudinal study, which employed the optimal design for the study of long-term treatment effects, did not find such treatment effects (Shaw et al., 2014), nor did previous analyses in our own sample (Greven et al., 2015; Schweren et al., 2015a).

The literature on long-term treatment effects in the human brain has, with few exceptions, focused on brain structure, while studies of acute stimulant effects focused on brain activation patterns. A single dose of methylphenidate has repeatedly been found to alter brain activation patterns in ADHD patients; case-control differences in blood-oxygen level dependent (BOLD)-response to cognitive/motivational tasks became smaller or disappeared when patients were on stimulant medication (Cubillo et al., 2012a). Little is known about whether acute functional changes translate into long-term functional changes as well. Adults with a history of untreated childhood ADHD showed blunted ventral-striatal activation compared to controls when exposed to emotional pictures, whereas adults with a history of ADHD who had received stimulant treatment during childhood did not (Schlochtermeier et al., 2011). During reward processing, the same group of treatment-naive adults showed lower insula activation compared to controls and childhood stimulant-treated adults (Stoy et al., 2011). These findings may suggest enduring functional therapeutic changes. In a meta-analysis of attention tasks, striatal activity was particularly reduced in studies including mostly stimulant-naive patients (Hart et al., 2013). Radioligand studies, however, have reported exacerbated rather than attenuated deficits in striatal dopamine neurotransmission after long-term stimulant treatment in adults with ADHD (Fusar-Poli et al., 2012; Ludolph et al.,

2008). Summarizing, stimulant treatment may be associated with persistent changes in brain activation patterns and/or dopamine metabolism, but the evidence is very limited and it remains unclear to what extent such changes may be therapeutic or disadvantageous.

The striatum is of particular interest when studying stimulant treatment effects in ADHD. Reduced striatal volumes (Frodl & Skokauskas, 2012; Nakao et al., 2011), lower striatal activity during reward anticipation and higher striatal activity during outcome of reward (Aarts et al., 2015; Paloyelis et al., 2012; Von Rhein et al., 2015b) have repeatedly been found in ADHD. Moreover, the striatum is rich in dopamine transporters, an important molecular target of stimulant treatment. Hence, long-term stimulant treatment effects may be expected to occur in the striatum. However, acute stimulant-induced changes in activation patterns have also been reported in supplementary motor areas (SMA), frontal cortex, anterior and posterior cingulate cortex, and precuneus cortex (e.g., Peterson et al., 2009, Prehn-Kristensen et al., 2011, Rubia et al., 2009).

We investigated associations between lifetime stimulant treatment history and neural activity during reward processing, using magnetic resonance (MRI) data from a large observational study. An innovative data-driven classification method was used to identify patient subgroups with distinct treatment trajectories. In our cohort, Groenman et al., (2015) found these trajectories to be clinically relevant for the development of substance use disorder (*unpublished data*). Moreover, treatment timing and dose have been found to moderate long-term stimulant treatment effects in the rat brain (e.g., van der Marel et al., 2014). In prior work, our group showed higher striatal BOLD-response to reward outcome in ADHD patients compared to controls (Von Rhein et al., 2015b). In the current study, we hypothesized that patients who had received more intense treatment would show reduced striatal BOLD-response (i.e., more similar to controls) to reward outcome compared to those who had received less intense treatment. Second, we hypothesized that between-group differences in other brain regions, if any, would show a similar pattern.

METHODS

Participants

Participants with ADHD were selected from the family-based IMAGE-NeuroIMAGE cohort (2009-2012) (von Rhein et al., 2015a). Children, adolescents, and young adults participated in diagnostic interviews, questionnaires, DNA collection, and an MRI session, taking place at two sites. Informed consent was signed by all participants ≥12 years old and all parents of participants <18 years old. The study

was approved by the local ethical committees of each participating site. Inclusion criteria were: IQ≥70, age 8-30 years, no diagnosis of classical autism, learning difficulties, brain disorders, or genetic disorders, and no contra-indication for MRI scanning. ADHD diagnosis (any type) was confirmed in accordance with the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; ADHD, 2013), operationalized as six or more symptoms on the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS; Kaufman et al., 1997) and t>63 on the Conners parent-, teacher-, and/or self-rated ADHD scales (Conners et al., 1998a, 1998b and 1999), rated while participants were off-medication. Five K-SADS symptoms were sufficient for diagnosis in participants age 16 or older, in line with DSM-5 revised criteria. The initial ADHD sample consisted of 269 participants. Functional MRI data were available for 124 patients (mean age=17.4 years, range 10-26 years).

Control participants were required to have no scores in the (sub)clinical range on any of the ADHD rating scales or interviews, no current or past psychiatric diagnosis or treatment, and no first-degree relative with ADHD. The initial control sample consisted of 187 participants. Functional MRI data was available for 97 controls (mean age=17.0 years, range 10-23 years).

Stimulant treatment

History of psychoactive treatment was assessed using pharmacy prescription records containing delivery date, substance name, dose, quantity, and frequency of use for each delivery between date-of-birth and date-of-scan. In addition, patients and parents participated in face-to-face semi-structured interviews to reconstruct lifetime treatment history. Self-report data was highly compatible with data derived from pharmacies (data not shown), with reliability estimates similar as those reported by Kuriyan et al. (2014) Self-report data was used only when pharmacy data was incomplete. Stimulant intake in mg (immediate- and extended-release methylphenidate preparations, and dexamphetamine preparations) was reconstructed for each day between date-of-birth and date-of-scan. Daily intake in mg was averaged for every month of the participant's life. Stimulant start age, stop age, and lifetime cumulative stimulant dose were calculated from this reconstruction. A smooth generalized additive model curve was fitted to each participant's reconstruction, allowing estimation of three additional treatment parameters that were more sensitive to noise, i.e., treatment duration (estimated stop age minus estimated start age), treatment variability (standard deviation of the fitted curve), and the lifetime maximum dose. Treatment duration and cumulative stimulant dose were adjusted for current age.

Community detection algorithm

The six stimulant treatment parameters (start age, stop age, total dose, estimated duration, estimated maximum daily dose, and estimated variability) were entered in an automated, optimization-based, weight-conserving community detection algorithm (Rubinov & Sporns, 2011). The algorithm categorizes participants into mutually exclusive communities (groups), segregating groups such that within-group positive/negative correlations are maximal while between-group correlations are minimal. The modularity statistic Q (range 0-1) quantifies the degree to which participants may be subdivided into clearly delineated groups. The algorithm terminates when Q no longer increases from one iteration to the next. Robustness of the optimal community structure was confirmed using non-parametric bootstrap procedures (Supplement S1).

TABLE 1. Stimulant treatment characteristics [mean(SD)] for participants in each treatment group.

	'early-and-intense' (41.3%)	'late-and-moderate' (35.7%)	'early-and-moderate' (7.4%)
Start age	6.89(1.29)	11.19(2.62)	8.25(1.48)
Stop age	15.65(2.97)	16.37(3.12)	11.37(2.01)
Treatment duration [a]	9.10(2.63)	3.99(2.47)	2.40(1.47)
Variability	320.76(459.27)	110.42(154.94)	66.43(88.48)
Cumulative dose in mg [a]	67480(55751)	38413(29986)	18437(17474)
Maximum daily dose in mg	47.37(23.36)	24.45(16.60)	17.53(12.61)

[a] Classification was based on an age-adjusted measure; the value reported here is calculated based on the non-smoothed trajectories.

The data-driven classification method produces more reliable results in larger samples, hence all participants with ADHD were included in this step (n=269). Stimulant-naive participants were a priori defined as a separate category (n=42, 15.1%). For stimulant-treated participants, the optimal solution yielded three treatment groups (Q=0.580; Table 1). The first group (n=111, 41.3%, 'early-and-intense') was characterized by early treatment onset, long duration, and a high maximum and total dose. The second group (n=96, 35.7%; 'late-and-moderate') was characterized by older age at treatment onset, shorter duration, and lower maximum and total dose. The third group (n=20, 7.4%; 'early-and-moderate') was characterized by early treatment onset, medium duration, and low maximum and total dose. As few

participants were classified to the 'early-and-moderate' group or were stimulant-naive, 'early-and-intense'-vs.-'late-and-moderate' was our primary contrast of interest. As shown in Table 1, the 'early-and-intense' and 'late-and-moderate' groups differed in stimulant start age, treatment duration, variability, maximum dose, and total dose, but not in stop age.

Reward task

A modified version of the monetary incentive delayed task was performed in the scanner (Von Rhein et al., 2015b). Participants were asked to respond as quickly as possible to a target by pressing a button. Before this target, a cue indicated the possibility of gaining a reward after a button press within a given time-window. Every trial ended with a feedback screen informing about the outcome of the current trial. Depending on the participant's performance, the response-window for a correct response was adapted in the next trial, resulting in an expected hit-rate of 33%. The experiment lasted 12 minutes, and a total of €5 could be gained. At the end of the experiment, the awarded money was paid to the participant. Compared with the original task, our version differed on two main aspects: hit-rate (33% versus 66%) and reward magnitude (€0.20 versus $5). The rationale behind these adaptations was firstly to increase the demands of the task with stronger task engagement as a result. Secondly, our adaptations aimed at meeting the practical constraints of our study. Considering that we limited ourselves to rewarded and neutral conditions, rewarding participants according to the original task parameters would have led to disproportionate monetary rewards (approximately €80), which was a concern for us and our ethical review board. Reaction time reward sensitivity was calculated as the mean reaction time across non-rewarded trials minus the mean reaction time across rewarded trials, with higher values indicating higher sensitivity to reward.

Functional MRI processing and analyses

Acquisition parameters, preprocessing steps, and first-level analyses were identical to those in our previous publication (Von Rhein et al., 2015b; Supplement S2). Second-level analyses for each task condition (reward anticipation and outcome) comprised both region of interest (ROI) and whole-brain analyses in FMRIB Software Library (FSL; Smith et al., 2004). First, main task effects were identified in a one-sample t-test, with scanner, age, gender, and three motion parameters as regressors of no interest. For the ROI analyses, average parameter estimate was extracted for each participant from the (warped) task-activated voxels within a binary mask of the striatum (caudate, putamen, and accumbens). In a linear mixed effect regression

model in SPSS (IBM, 2013), striatal activation was predicted from treatment group (primary contrast: 'early-and-intense'-vs.-'late-and-moderate'; secondary contrasts: 'stimulant-naive'-vs.-'early-and-intense', 'stimulant-naive'-vs.-'late-and-moderate', 'stimulant-naive'-vs.-'early-and-moderate', 'early-and-moderate'-vs.-'early-and-intense', 'early-and-moderate'-vs.-'late-and-moderate'). Gender, scanner, age, and age^2 (to account for non-linear developmental trajectories of reward-related striatal activation) were added as covariates, along with a random intercept per family to account for relatedness within the sample. Alpha was adjusted for analyzing one primary and five secondary contrasts and two task conditions ($\alpha=0.05/6/2=0.004$). Normalized first-level b-maps were also entered into whole-brain second-level mixed effect analyses. Treatment group was entered as a predictor along with scanner, gender, age, and three movement parameters ($Z_{VOXEL}>2.3$; $\alpha_{CLUSTER}=0.004$).

Structural MR images were also acquired, to assess structural correlates of long-term functional changes, if any (Supplement S3).

Follow-up and sensitivity analyses

For each whole-brain significant cluster, average parameter estimate was extracted per participant for follow-up analyses in SPSS. Treatment groups were data-driven, hence not matched with regard to clinical and demographic variables. Potential confounders other than age and gender (i.e., IQ, SES, ADHD symptoms, ADHD-type, comorbidity, and history of non-stimulant psychoactive medication) were added to the model. Moreover, analyses were repeated within one-to-one age-, gender-, and ADHD symptom count-matched subsamples (n=25 per group).

To exclude acute withdrawal/rebound effects, each significant effect was re-estimated separately for participants who were on active stimulant treatment within two weeks prior to scanning and those who had ceased treatment more than two weeks prior to scanning.

Main reward task effects and case-control differences in the current cohort have previously been reported (Von Rhein et al., 2015b), hence are not addressed here. For reference only, the control sample mean for each outcome measure was estimated in a covariate-only model.

RESULTS

Sample characteristics

The ADHD sample consisted of 83 males (66.9%) and 41 females (33.1%), with an average age of 17.4 years (SD=3.0, range 10-26 years; Table 2). Of those, 51

TABLE 2. Characteristics of the 'early-and-intense' and 'late-and-moderate' treatment groups [N(%), unless otherwise specified].

	ADHD N=124	E&I N=51	L&M N=49	
Male	83(66.9)	46(90.2)	26(53.1)	Chi^2=17.1*
Site Nijmegen	79(63.7)	32(62.7)	36(73.5)	Chi^2=1.3
Age *M(SD)*	17.4(3.0)	17.1(2.4)	18.1(3.0)	F=3.2
IQ *M(SD)*[a]	99.2(15.0)	98.6(14.4)	100.2(14.6)	F=0.3
Current stimulant users	46(42.2)	30(58.8)	13(26.5)	Chi^2=10.6*
Symptoms of inattention *M(SD)*	7.2(1.8)	7.8(1.3)	6.6(2.0)	F=11.5*
Symptoms of hyperactivity/impulsivity *M(SD)*	6.0(2.4)	6.7(2.3)	5.7(2.2)	F=5.4
ADHD-type				
Inattentive	56(45.2)	18(35.3)	24(49.0)	Chi^2=1.9
Hyperactive/impulsive	17(13.7)	6(11.8)	9(18.4)	Chi^2=0.9
Combined	51(41.1)	27(52.9)	16(32.7)	Chi^2=4.2
Comorbidity				
ODD-CD	30(24.2)	17(33.3)	8(16.3)	Chi^2=3.9
Tic disorder	1(0.8)	1(2.0)	0(0.0)	Chi^2=1.0
Anxiety/depression	3(2.4)	1(2.0)	1(2.0)	Chi^2<0.1
Substance use disorder[b]	24(19.4)	8(15.7)	14(28.6)	Chi^2=2.4
Non-stimulant medication				
Atomoxetine	19(15.3)	10(19.6)	8(16.3)	Chi^2=0.2
Antipsychotics	23(18.5)	16(31.4)	5(10.2)	Chi^2=6.8
Anxiolytics	8(6.5)	3(6.1)	3(6.1)	Chi^2<0.1
Antidepressants	8(6.5)	4(7.8)	2(4.1)	Chi^2=0.6

E&I, early-and-intense; L&M, late-and-moderate; ODD-CD, oppositional defiant disorder-conduct disorder; [a] estimated based on the 'vocabulary' and 'block design' subtests of the Wechsler intelligence scales for children/adults. [b] assessed approximately two years prior to participation in the current study. *p<0.004.

participants were assigned to the 'early-and-intense' treatment group (46.8%), and 49 to the 'late-and-moderate' group (45.0%). Compared to the 'late-and-moderate' treatment group, the 'early-and-intense' group contained more males and more participants on active stimulant treatment, and had more inattention problems. The two groups did not differ with regard to age, socio-economic status, IQ, ADHD-type, hyperactivity/impulsivity symptoms, comorbidity, or history of non-stimulant medication. The control sample (n=97; age M=17.0 years, SD=2.9, range 10-23 years) contained fewer males compared to the ADHD sample (44.3% vs. 66.9%; p=0.001). For the stimulant-naive (n=15) and 'early-and-moderate' (n=9) groups, see Supplement S4.

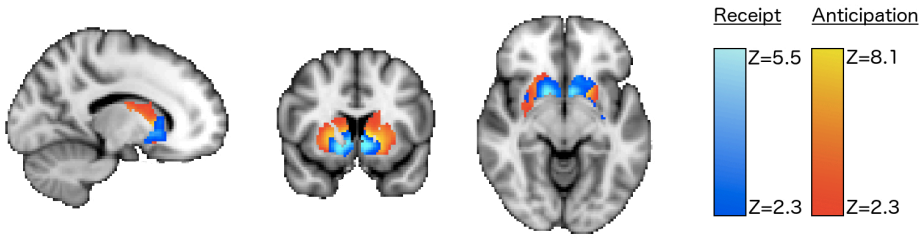


FIGURE 1. Striatal activation during reward anticipation in yellow-red, and during reward outcome in blue-light blue, across all participants.

Reward processing

The striatum was activated by both task conditions (Figure 1). There were no differences in striatal BOLD-response between the 'early-and-intense' and 'late-and-moderate' treatment groups during reward anticipation ($M_{EARLY\&INTENSE}$=360.7, $M_{LATE\&MODERATE}$=394.8, $M_{CONTROL}$=299.7, p=0.784), nor during reward outcome ($M_{EARLY\&INTENSE}$=362.1, $M_{LATE\&MODERATE}$=677.5, $M_{CONTROL}$=414.9; p=0.180).

Whole-brain functional MRI analyses did not yield any clusters of significant difference between the 'early-and-intense' and the 'late-and-moderate' groups during reward anticipation. In the reward outcome condition, the 'late-and-moderate' group showed lower activity compared to the 'early-and-intense' group in a cluster located in the SMA, extending into the dorsal anterior cingulate cortex (dACC) and paracingulate gyrus (Figure 2; $M_{EARLY\&INTENSE}$=635.1, $M_{LATE\&MODERATE}$=-813.9, $M_{CONTROL}$=35.5, cluster size=8696 mm^3, B=-1449.0, $p_{CLUSTER}$<0.001). Gender (B=964.6, p=0.014), scanner (B=179.0, p=0.604), age (B=-285.8, p=0.087), and age^2 (B=153.8, p=0.087) were not associated with activation in this cluster, nor were any of the additional covariates when added to the model while the effect of treatment history remained unchanged. Moreover, the pattern was consistently observed in past

users ($M_{EARLY\&INTENSE}$=374.9, $M_{LATE\&MODERATE}$=-687.3) and current users ($M_{EARLY\&INTENSE}$=785.0, $M_{LATE\&MODERATE}$=-1323.6), and within the age-, gender-, and symptom-matched subsamples ($M_{EARLY\&INTENSE}$=721.5, $M_{LATE\&MODERATE}$=-395.5).

There was no behavioral (i.e., reaction time) difference in reward sensitivity between the 'early-and-intense' and 'late-and-moderate' groups ($M_{EARLY\&INTENSE}$= 35.0ms, $M_{LATE\&MODERATE}$=29.4ms, $M_{CONTROL}$=25.7ms, p=0.559). Moreover, reaction time reward sensitivity was not associated with striatal activity during reward anticipation (Pearson r=0.173, p=0.055) or reward outcome (Pearson r=0.014, p=0.879), nor with activity within the SMA/dACC cluster (Pearson r=0.177, p=0.050).

There were no structural brain differences between the two groups. For findings involving the 'early-and-moderate' and stimulant-naive groups, see supplement S4.

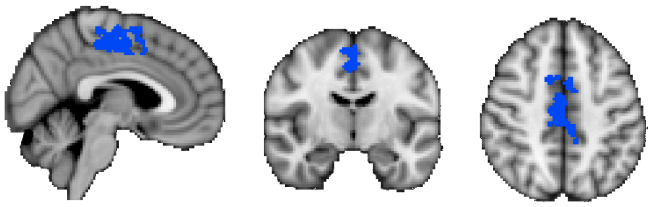


FIGURE 2. Higher activation in the 'early-and-intense' group compared to the 'late-and-moderate' group during reward outcome.

DISCUSSION

In a large sample of children, adolescents and young adults with ADHD, we investigated whether characteristics of stimulant treatment history were associated with brain activation patterns during reward processing while off medication. Stimulant treatment history was not associated with BOLD-response to reward anticipation or outcome in the striatum. In the SMA/dACC, individuals with a history of moderate treatment showed lower activity during reward outcome compared to those with a history of intense treatment. While activity in the moderately treated group was reduced compared to controls, activity in the intensely treated group was higher compared to controls. Our findings thus suggest compensatory SMA/dACC recruitment in individuals with a history of intense stimulant treatment. The effect is likely driven by treatment duration and dose rather than recency of treatment discontinuation, since stop age did not differ between the two groups.

Higher striatal BOLD-response to reward outcome has consistently been reported in ADHD (e.g., Paloyelis et al., 2012; Von Rhein et al., 2015b). As such changes have been shown to disappear after stimulant administration (Aarts et al.,

2015; Knutson et al., 2004), we had hypothesized that participants with a history of intense treatment would show lower striatal BOLD-response to reward outcome compared to those with a history of less intense treatment. We found no evidence for such an effect. Moreover, there was no association between treatment history and striatal activity during reward anticipation. Our findings may indicate that the acute changes in striatal activity in response to stimulants do not translate into lasting functional changes in this region during reward processing. This finding is consistent with Stoy et al. (2011) who, in a small adult sample, also reported no changes in striatal activation during reward outcome after childhood stimulant treatment.

We found a large cluster of lower activity during reward outcome in the moderately treated subgroup compared to the intensely treated subgroup, located in the bilateral SMA and dACC, extending into the precuneus and posterior cingulate cortex. Dorsal and mid-cingulate regions project to the ventral striatum, and are important for monitoring incentive-based behavioral responses (Haber & Knutson, 2010; Shima et al., 1998). Hypo-activation has previously been reported in medication-naive ADHD patients during reward outcome (Carmona et al., 2012). Acute stimulant effects in the SMA/dACC during reward processing have been reported as well (Aarts et al., 2015), although most fMRI studies of reward reported no acute stimulant effects in this region (e.g., Knutson et al., 2004; Rubia et al., 2009).

Lower activity in the SMA/dACC in ADHD patients has also been associated with cognitive processes other than reward processing. Higher SMA/dACC activation may represent recruitment of a cognitive process enhancing feedback-based decision-making, even when a motor response is not required (Ullsperger & Von Cramon, 2003; Vassena et al., 2015), as was the case in the reward outcome phase of our task. ADHD patients have shown lower SMA activity when selection of a non-habitual response was required (Hart et al., 2013; Suskauer et al., 2008). Higher SMA/dACC, PCC, and precuneus activity has been reported after a single dose of stimulants during tasks requiring feedback-based modulation of motor responses (Bush et al., 2008; Cubillo et al., 2012b; Pliszka et al., 2007), but acute effects in the opposite direction have also been reported (Ivanov et al., 2014; Rubia et al., 2014). Enhanced cognitive decision-making upon reward in intensely treated individuals is consistent with the lower rate of SUD in this group (Groenman, *unpublished data,* 2015), although the difference in SUD rate in the current (smaller) fMRI sample was not significant. Summarizing, higher SMA/dACC activity may indicate enhanced cognitive decision-making following reward after early and high-dose stimulant treatment. Note that this proposition is not supported in behavioral data, as our paradigm required no response following reward outcome.

Alternatively, higher SMA/dACC activity may represent increased salience network activity, enhancing attention allocation to emotional, rewarding, or

surprising events (Menon, 2015). Stimulant-induced improvement in cognitive performance has been shown to be mediated by enhanced salience (Jolles et al., 2011; Ter Huurne et al., 2015). Stimulant treatment history may be associated with greater focus on the task. Yet, increased task focus may be expected to occur throughout the task as opposed to during the outcome phase only, and may result in improved task performance which we did not observe. Finally, higher SMA/dACC activity may entail enhanced 'readiness to act' upon reward outcome, as the SMA is embedded in the task-positive motor network (Volkow et al., 2004). However, we found no association between SMA/dACC activation and reaction times.

The current study has several strengths. First, only a handful of prior studies investigated functional rather than anatomical long-term neural changes in relation to stimulant treatment in ADHD. Of those, the current sample is by far the largest. Second, the data-driven classification of participants with ADHD based on multiple treatment characteristics is novel and clinically relevant. The current study has limitations as well. Long-term treatment effects can only be studied observationally. Although findings have been statistically adjusted for group differences, confounding by indication could not be excluded. Moreover, few participants were stimulant-naïve (in accordance with high prescription rates), and data-driven classification of stimulant-treated participants yielded unbalanced groups. This allowed powerful analysis of participants in the two largest groups, but restricted analyses of stimulant-naive participants and those with early-and-moderate treatment. Finally, no data was collected regarding behavioral treatment, which, according to guidelines, should be offered in conjunction with pharmacological treatment; hence, pharmacological and behavioral treatment effects cannot be distinguished in our study. The recruitment of compensatory cognitive control areas may reflect the application of cognitive strategies learned during behavioral treatment.

We conclude that ADHD patients with a history of early-onset high-dose stimulant treatment showed more SMA/dACC activation during reward outcome, compared to those with a history of late-onset moderate-dose stimulant treatment. Higher SMA/dACC activity may represent a compensatory mechanism of enhanced higher-level processing of reward information in the intensely treated group. Stimulant treatment history was not associated with striatal BOLD-response to reward processing. Understanding long-term risk and benefits of stimulant treatment could be further enhanced by evaluating functional rather than neuroanatomical brain changes in future studies.

SUPPLEMENTAL INFORMATION

S1. Robustness of the community detection algorithm

The community detection algorithm was rerun one thousand times, while randomly selecting 227 times one participant with replacement for each run (bootstrapping). Participants could be duplicated within a bootstrapped sample. The algorithm produced a three-class solution in 793 runs (79.3%), with an average modularity Q of 0.58 (SD=0.02; confidence interval=0.579-0.582). When a three-class solution was found, the largest class contained on average 46.0% of participants (SD=3.5%; confidence interval=45.7% - 46.3%), while the second largest class contained on average 40.9% of participants (SD=3.0%; confidence interval=40.7% - 41.1%) and the smallest class contained on average 13.2% of participants (SD=5.4%; confidence interval=12.8% - 13.5%). Thus, the average distribution of classes across 1000 runs strongly resembled the solution reported in the paper. Moreover, the 95% confidence intervals are narrow and standard deviations are low, indicating that the percentage of participants per class tend to be stable across runs. As expected, the percentage of participants in the smallest class is most susceptible to random variations (SD=5.4%).

The algorithm produced a four-class solution in 133 runs (13.1%). The four classes contained on average 43.0% (SD=3.4), 38.7% (SD=3.3), 11.6% (SD=3.7) and 6.6% (SD=3.2) of participants, respectively. In 76 runs, the algorithm produced a two-class solution (7.6%), with classes containing on average 54.1% (SD=2.9) and 45.9% (SD=2.9) of participants, respectively. Thus, the three-class solution reported in the paper resembles the distribution of participants across the three largest classes of the four-class solution, as well as the distribution of participants in the two-class solution. This underlines the stability of the three-class model.

S2. Functional MRI - acquisition and preprocessing

MRI data was acquired on two Siemens 1.5 Tesla scanners (Erlangen, Germany) with matched head coils and acquisition parameters. Participants were randomly assigned a combination of three or four functional acquisitions (a diffusion weighted scan, resting-state scan, reward task, working memory task, and/or response inhibition task). Reward task functional MRI data was thus available for a random subset of participants (n_{ADHD}=124, n_{HC}=97). Whole brain functional imaging was performed using a gradient-echo echo-planar scanning (EPI) sequence (37 axial slices, TR=2340ms, TE=40ms, voxel size=3.5x3.5x3.0mm, inter-slice gap=0.5mm, FOV=224mm, FA=90°). Participants with more than three head movements of ≥4mm during the task were excluded.

Functional MRI preprocessing steps included spatial realigning, nuisance regression, spatial smoothing at FWHM=6mm. First-level statistical parametric maps (b-maps) were estimated for each participant, including 6 regressors of interest (onset times of non-rewarded and rewarded cues, hits, and misses), and 6 regressors of no interest (onset times of non-rewarded and rewarded targets, onset times of targets, cues, and outcomes followed by incorrect responses, and a motion regressor identifying and excluding events affected by excessive movement). Participants with less than five occurrences of one or more event types (rewarded hits, rewarded misses, non-rewarded hits, or non-rewarded misses) were excluded. All regressors and their temporal derivatives were convolved with a canonical hemodynamic response function. For reward anticipation, response maps for rewarded cues were contrasted with response maps for non-rewarded cues. Activation during reward outcome was assessed by the interaction of accuracy (hits versus misses) and reward (rewarded versus non-rewarded trials). First-level b-maps were registered using non-linear transforms to a study-specific template.

S3. Structural MRI – acquisition, processing, and analyses

The MRI session included at least one T1-weighted structural acquisition (3D MP-RAGE; 176 sagittal slices, TR=2730ms, TE=2.95ms, voxel size=1x1x1mm, FOV=256mm, FA=7°, parallel imaging by generalized auto-calibrating partially parallel acquisition [GRAPPA]). For each participant, the structural acquisition of highest quality was selected by visual inspection, accepting only scans with no or minimal distortions. Structural MRI analyses were performed in the initial samples (n_{ADHD}=269; n_{HC}=187), to increase power.

Total striatal volume (sum of left and right putamen, caudate, and accumbens) was calculated using FSL FIRST with default settings (Patenaude et al., 2011). ROI analyses were performed in SPSS, predicting striatal volume from treatment group, with covariates gender, scanner, age, age^2, total brain volume (TBV, calculated with the VBM8 toolbox in SPM; Ashburner & Friston, 2005), and a random intercept per family. Whole-brain structural analyses included volumetric analyses of additional subcortical structures (globus pallidus, amygdala, hippocampus, and thalamus; sum of left and right hemisphere; α=0.008/4=0.002) and vertex-wise analysis of cortical thickness and surface area. Freesurfer with default settings was used to reconstruct the cortical surface of each participant (Dale et al., 1999; Fischl et al., 1999; Fischl & Dale, 2000).

Statistical maps were computed for the 'early-and-intense' *vs.* 'late-and-moderate' contrast and the five secondary contrasts, with age, gender, and scanner as covariates, and age^2 as an additional per-vertex-regressor. Normalized statistical maps were thresholded per vertex ($Z > 2.3$). Cluster-level thresholding was based on Monte Carlo simulation testing, with $\alpha_{CLUSTER}$ adjusted for testing six contrasts and two hemispheres (α=0.004). For each significant cluster, a random intercept per family was added to the model in SPSS.

S4 - Secondary contrasts

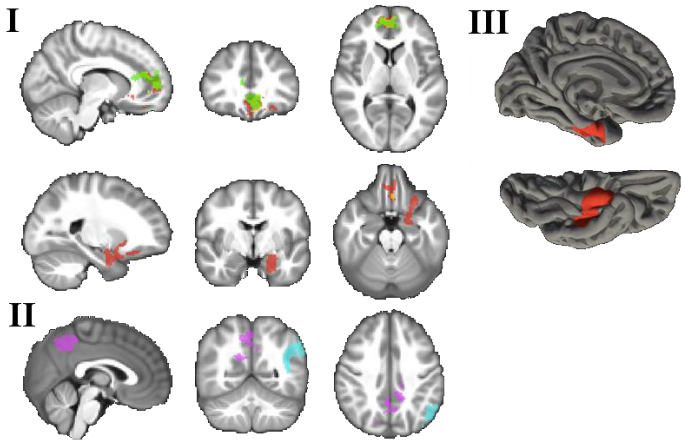


FIGURE S2.1. Whole-brain functional and structural MRI results for secondary contrasts. Left panel: significant between-group differences in BOLD-response during reward anticipation. I: 'early and moderate' < stimulant naïve in red, 'early and moderate' < 'late and moderate' in yellow, 'early and moderate' < 'early and intense' in green. II: 'early and moderate' < stimulant naïve in purple, 'late and moderate' < stimulant naïve in blue. Right panel (III): cluster of decreased cortical surface area in the 'early and moderate' compared to the 'early and intense' treatment group. Follow-up analyses showed that surface area in this cluster was also associated with gender (data not shown).

TABLE S4.1. Sample characteristics and significant confounders in the five secondary contrasts.

	L&M N=49	E&I N=51	E&M N=9	NAIVE N=15
Male *N(%)* [I, II]	26 (53.1)	46 (90.2)	4 (44.4)	7 (46.7)
Nijmegen *N(%)* [III]	36 (73.5)	32 (62.7)	6 (66.7)	5 (33.3)
Age in years *M(SD)* [IV]	18.1 (3.0)	17.1 (2.4)	14.9 (2.5)	17.5 (4.0)
Estimated IQ *M(SD)* [a]	100.2 (14.6)	98.6 (14.4)	101.1 (15.8)	97.3 (18.8)
SES *M(SD)*	11.4 (2.2)	11.6 (2.1)	10.6 (1.7)	12.1 (2.5)
Current stimulant users *N(%)*	13 (26.5)	30 (58.8)	3 (33.3)	0 (0.0)
Inattention sympt. *M(SD)*	6.6 (2.0)	7.8 (1.3)	7.8 (0.8)	6.7 (2.1)
Hyperactive/impulsive sympt. *M(SD)* [II]	5.7 (2.2)	6.7 (2.3)	5.4 (2.0)	4.5 (3.0)
ADHD type				
Inattentive *N(%)*	24 (49.0)	18 (35.3)	6 (66.7)	8 (53.3)
Hyperactive/impulsive *N(%)*	9 (18.4)	6 (11.8)	0 (0.0)	2 (13.3)
Combined *N(%)*	16 (32.7)	27 (52.9)	3 (33.3)	5 (33.3)
Comorbidity				
ODD-CD *N(%)*	8 (16.3)	17 (33.3)	2 (22.2)	3 (20.0)
Tic disorder *N(%)*	0 (0.0)	1 (2.0)	0 (0.0)	0 (0.0)
Anxiety/depression *N(%)*	1 (2.0)	1 (2.0)	1 (11.1)	0 (0.0)
Substance use disorder *N(%)* [b]	14 (28.6)	8 (15.7)	1 (11.1)	1 (6.7)
Non-stimulant treatment				
Atomoxetine *N(%)*	8 (16.3)	10 (19.6)	1 (11.1)	0 (0.0)
Atypical antipsychotics *N(%)*	5 (10.2)	16 (31.4)	2 (22.2)	0 (0.0)
Anxiolytics *N(%)*	3 (6.1)	3 (6.1)	1 (11.1)	1 (8.3)
Antidepressants *N(%)*	2 (4.1)	4 (7.8)	0 (0.0)	1 (8.3)

L&M, late & moderate; E&I, early & intense; E&M, early and moderate; ODD-CD, oppositional defiant disorder-conduct disorder; SES, socio-economic status; [a] estimated based on the 'vocabulary' and 'block design' subtests of the Wechsler intelligence scales for children/adults. [b] assessed approximately two years prior to participation in the current study. Significant differences between groups ($\alpha<0.004$): [I] early & intense *vs.* early & moderate; [II] naïve *vs.* early & intense; [III] naïve *vs.* late & moderate; [IV] early & moderate *vs.* late & moderate.

TABLE S4.2. Mean reaction time reward sensitivity and striatal region-of-interest results for each secondary contrast.

	RT-RS in ms.		$BOLD_{ANTICIPATION}$		$BOLD_{OUTCOME}$		Volume (in mL)	
	M		M		M		M	
NAÏVE	35.5		540.7		671.8		19.8	
L&M	29.4		394.8		362.1		20.1	
E&M	45.1		180.2		1035.7		20.1	
E&I	35.0		360.7		677.5		20.3	
	B	p	B	p	B	p	B	p
NAIVE *vs.* L&M	6.1	0.647	145.8	0.403	309.7	0.345	-0.3	0.253
NAIVE *vs.* E&M	-9.6	0.606	360.4	0.140	-363.8	0.427	-0.3	0.511
NAIVE *vs.* E&I	0.5	0.970	180.0	0.332	-5.7	0.987	-0.5	0.098
E&M *vs.* L&M	15.6	0.331	-214.6	0.310	673.6	0.091	-0.1	0.893
E&M *vs.* E&I	10.1	0.535	-180.4	0.398	358.1	0.372	-0.2	0.567

RT-RS, reaction time reward sensitivity; L&M, late & moderate; E&M, early & moderate; E&I, early & intense; M, estimated marginal means in a model with covariates gender, scanner, age, and age^2 (and total brain volume for volumetric analyses).

TABLE S4.3. Whole-brain functional and structural MRI results for each secondary contrast

Condition/measure	Brain region		M_{NAIVE}	$M_{L\&M}$	$M_{E\&M}$	$M_{E\&I}$	Sign. contrast	B	Cluster size	$p_{CLUSTER}$
BOLD $_{ANTICIPATION}$	ACC [a]	B	172.3	65.3	-1510.0	-214.6	E&M < L&M	-1575.3	6408 mm^3	0.00399
	ACC [a]	B	123.6	-244.0	-1452.7	-101.1	E&M < E&I	-1351.6	7696 mm^3	0.00157
	ACC, OFC [a]	B	322.7	-158.2	-1196.9	-224.7	E&M < NAIVE	-1519.6	9032 mm^3	0.00074
	Precuneus	B	255.6	-366.1	-1294.2	-102.0	E&M < NAÏVE	-1549.8	16400 mm^3	0.00001
	Sup. parietal	L	54.2	-1094.1	-1061.3	-539.8	L&M < NAÏVE	-1148.3	7128 mm^3	0.00007
BOLD $_{OUTCOME}$	*None*		-	-	-	-	*None*	-	-	-
Cortical thichness	*None*		-	-	-	-	*None*	-	-	-
Surface area	Inf. temporal	L	1300.9	1385.8	1171.0	1395.4	E&M < E&I	-224.4	995 mm^2	0.00340
Volume	Hippocampus	B	7.9	7.8	7.6	7.9	*None*	-	-	-
	Amygdala	B	2.7	2.6	2.7	2.7	*None*	-	-	-
	Thalamus	B	16.7	16.7	16.5	17.0	*None*	-	-	-
	Gl. Pallidus	B	3.7	3.7	3.8	3.8	*None*	-	-	-

L&M, late & moderate; E&M, early & moderate; E&I, early & intense; OFC, orbitofrontal cortex; L, left hemisphere; R, right hemisphere; B, bilateral; M, estimated marginal means in a model with covariates gender, scanner, age, age^2 (and total brain volume for volumetric analyses). [a] clusters partially overlap