Functional Neuroimaging Evidence for Distinct Neurobiological Pathways in Attention-Deficit/Hyperactivity Disorder

Supplemental Information

Clinical Assessment Measures

In addition to the KSADS-PL semi-structured interview that yielded both Axis I DSM-IV diagnoses and ADHD symptom counts for analysis, the clinical assessment included parent- and participant-reported ADHD-related problem behavior on the Brown ADD Scales (1). General behavioral problems of clinical relevance were assessed using the Child Behavior Checklist (CBCL; (2)). Specific depressive and anxiety symptom severity were quantified using the Beck Depression Inventory II (BDI-II; (3)) and Multidimensional Anxiety Scale for Children (MASC; (4)). Tobacco use was measured by the Fagerstrom Test for Nicotine Dependence (5) and illicit substance use by the Adolescent Alcohol/Drug Use Scale (AADIS; (6)). Familial ADHD history was assessed in first-degree relatives by parent interview using the Family History Screener (7).

Neuropsychological Tests

The cognitive test battery was constructed to fully assess two neurocognitive domains that have reliably proven to be impaired in ADHD – motor inhibition tasks that emphasize executive control (8, 9), and several that assess different aspects of choice impulsivity, i.e., a preference for smaller, immediate rewards (10).

The executive/inhibition tests included Conner's Continuous Performance Test-II (CPT-II (11)). The CPT-II is a visual speeded response test that measures sustained attention, impulsive responding, motor response variability, response style and several other useful metrics.

Participants press the space bar of a computer keyboard when any letter except "X" appears on the screen (10% of 360 trials). The 250 msec trials are presented in 18 blocks of 20 trials, requiring 14 minutes of test administration time. The primary dependent measure examined was the number of commission errors. The Stop Signal Reaction Time Test (SSRT) is a classic assessment of the "horse race" motor inhibition model (12) in which two signals to execute or countermand a motor response compete. In this version (13), the auditory cue stop signal was presented on 25% of trials and the stop signal delays were automatically adjusted trial-by-trial until participants were able to withhold approximately 50% of trials. The stop signal reaction time (SSRT) was calculated by subtracting the final mean tone delay from the mean go reaction time (14). The Immediate and Delayed Memory Tasks (IMT/DMT) are another speeded visual CPT variant (15), but use different stimuli and conditions than Conner's CPT-II (e.g., the use of visually similar "lure" trials to provoke an impulsive response). One key difference is that the DMT includes a delay in which the stimuli must be retained in working memory in order to successfully guide performance on the subsequent trial. An impulsivity dependent measure was constructed for both the IMT and DMT (number of commission errors / number of catch trials) / (number of correct detections / number of target trials). The Matching Familiar Figures Test (MFFT) measures participants' impulsive non-reflective style. Like the original non-computerized version (16), participants were shown one standard picture and had to choose among six seemingly-identical variants to identify perfect matches. The primary dependent variable was the number of error responses, with higher error counts indicating a more impulsive, non-reflective style. Collectively, these tasks measured motor response inhibition as conceptualized in several different theoretical models and under different information processing contexts.

The reward tasks were chosen to measure impulsive choice tendencies. The Delay Discounting Questionnaire (DDQ; (17)) measures discounting or devaluation of hypothetical rewards as a function of delay. Participants answered 100 questions like: "Would you prefer $10.00 in 30 days or $2.00 today? The primary dependent measure was calculated as area-under-the-curve (AUC) under the empirically calculated hyperbolic discounting function (18). The Experiential Discounting Task (EDT; (19)) also assesses discounting, but requires participants to make choices for monetary reward that result in immediate consequences. Despite theoretical similarities, the EDT is thought to capture a somewhat different construct of discounting compared to the DDQ (20). For direct comparability of measures, the EDT also used AUC for its primary dependent measure. The Single Key Impulsivity Paradigm (SKIP; (21)) is an unstructured assessment where participants respond as often as they choose to obtain a reward. Longer delays between responses result in larger reward. The average inter-response interval is considered an assessment of an impulsive choice-based response tendency. Of all the reward-oriented tasks in the battery, it probably can be conceptualized as best measuring ADHD "delay aversion" tendencies. The primary dependent measure was the average inter-trial interval of free responses collected over a 20 minute assessment.

Taxometric Analysis Data Preparation Using Principal Component Analysis Data Reduction

A non-significant Little's MCAR test indicated that any missing values due to participants' occasional failure to complete every cognitive measure in the test battery were random. Scores were then inspected to identify any implausible or un-useable data, as can potentially occur on some of the tests chosen for the study. Values were recoded as missing if SSRT Z scores > 1.96, if IMT/DMT performance had no errors, if there were fewer than 5 SKIP

responses overall, or if a hyperbolic curve could not be fit and AUC=1 for any of the three phases of the DDQ or EDT. After excluding missing or non-useable data, ADHD participants had 89.7% and non-ADHD 90.8% of their test battery scores available for subsequent analyses. To maximize use of the available data, single imputation via multiple regression was used separately for each diagnostic study group to replace missing values in all cognitive test scores examined. Next, multiple regression was used to regress out any linear effects of age and WASI-estimated Full Scale IQ so that pathway analyses would not be driven by these effects. The residuals were saved as *z* scores and reoriented so positive values reflected "better" test performance for all measures to ease principal component analysis (PCA) interpretation.

To prepare data for taxometric analyses, PCA with Varimax rotation and Kaiser normalization was used to reduce the battery of neurocognitive scores into a smaller number of factors. A key question was whether to reduce data from the entire sample simultaneously, or just ADHD adolescents. The former approach presumes both groups show similar relationships among different tests. To address this assumption's validity, we compared the results of PCAs done separately for ADHD and non-ADHD subsamples. Both solutions had three factors with eigenvalues >1 that explained 54.3% (ADHD) or 55.1% (non-ADHD) of the variance. Though there were some commonalities between the two PCA solutions, the ADHD study group's factor structure deviated enough from non-ADHD to conclude that their abnormalities could not be considered as extremes on a normal distribution (Table S2). As judged by loading coefficients >.30, the first factor of each PCA solution primarily reflected executive control over motor responses for both groups, while the second two factors reflected reward-related test profiles. There were notable differences for specific tests, e.g., SSRT score prominently loaded onto the first PCA factor for non-ADHD, but co-varied with SKIP performance in ADHD. For the

second and third factors, there were study group differences in the role of the DDQ or a measure of impulsive cognitive style from a computerized version of the MFFT. Interestingly, in ADHD performance on the DDQ and EDT covaried. In contrast, for non-ADHD, these tests loaded onto different factors. The latter reflects findings that these measures assess somewhat different constructs in non-clinical samples (20). But in ADHD, it seems that DDQ and EDT performance is more closely related.

Taxometric Analyses

Each indicator used for taxometric analysis adequately and uniquely represented the constructs of interest, as assured by the removal of IQ- and age-related nuisance covariance from the neurocognitive data and PCA to ensure minimal correlation among factors. Taxometric analysis was done only for ADHD participants. Taxometric results were confirmed using a "multiple methods" approach in which the curve that depicts the latent structure in neurocognitive data was tested with two different taxometric measures to test consistency (MAMBAC (22) and MAXSLOPE (23)). Although each taxometric measure is based on the same formal definition of the taxonic latent structure captured in the general covariance mixture theorem, they use unique mathematical operations to test model fit hypotheses. As such, they provide complementary pieces of evidence for the nature of the taxa and participant assignment, as recommended for taxometric analyses (24). Figure S1 displays the average fit curves for the procedures for the taxometric analysis that divided the data into an n=46 taxon and n=71 complement (i.e., the first subgrouping "cut" where cognitively unimpaired ADHD were identified). These curves were super-imposed on comparison data generated from an iterative simulation technique that reproduces the distributional and correlational properties of the

neurocognitive data in order to visualize what the curves should resemble if the latent structure is either categorical or dimensional. Both clearly depict a categorical structure with CCFI values of .638 and .621, and estimated taxon base rates *P* of .398 and .412, respectively. Another taxometric analysis performed on the *n*=71 complement provided similar results dividing it into two subgroups (MAMBAC/MAXSLOPE CCFI=.580/.682).

For comparison purposes, the same analyses were run using non-ADHD PCA-derived indicator data. In contrast to the categorical model fit found for ADHD, these CCFI and simulation results suggested a smoothly dimensional latent structure in non-ADHD.

Neuropsychological Test Scores Not Used for Taxometric Analysis

In analyses that covaried for sex, we examined other scores produced by our neuropsychological test battery not used in taxometric classification because they were not directly theoretically related to the dual-pathway model (Supplementary Table S3). Briefly, all three ADHD subgroups showed differences from vs. non-ADHD differences in many tests of sustained attention, response speed, response variability, signal detection, and response bias. Also, response speed and signal detection scores differed among the three ADHD subgroups. To further test whether these differences were independent from the executive and reward-related pathways, we re-did these analyses, covarying for the PCA factor scores used for taxometric analyses. All ADHD subgroups' differences on these other cognitive tests were no longer significant, suggesting the pathways identified during taxometric analyses also drove the subgroup differences in response speed and signal detection.

MRI Scanning Procedures

All MRI data were collected using a 3.0 Tesla Siemens Allegra MRI scanner at the Olin Neuropsychiatry Research Center, Institute of Living at Hartford Hospital. Functional MRI data for this task was acquired using a gradient-echo echo-planar imaging (EPI) pulse sequence: repetition time (TR)=1500 msec, echo time (TE)=28 msec, flip angle=65°, field of view=24×24 cm, acquisition matrix=64×64, A>>P phase encoding, voxel size=3.4×3.4 mm, slice thickness=5 mm, number of slices=29 (acquired sequentially). Gradient echo fieldmaps: TR=580 msec, TE=7 msec, flip angle=90°, matrix=128×128, A>>P phase encoding, 3 mm slice thickness. MPRAGE T1-weighted images of brain structure: TR=2500 msec, TE=2.74 msec, flip angle=8°, matrix=256×208, 1 mm slice thickness.

During the Go/NoGo (GNG) fMRI task, participants were instructed to make a speeded button press with their right index finger to rapidly-presented visual 'X' (Go) stimuli (P = 0.85), but to withhold response to infrequent, interspersed 'K' (No-Go) stimuli (P = 0.15). All stimuli were presented for 50 msec. Prior to beginning the task, each participant performed a practice trial to ensure understanding of the instructions. Hits and errors were defined as a response occurring within 1,000 msec of an 'X' or 'K' trial, respectively. The two conditions of interest were false alarm responses (i.e., hitting the button incorrectly for a 'K' stimulus) and correct rejects (i.e., accurately withholding prepotent response to 'K' stimuli). The 'X' stimuli were so frequent that their hemodynamic responses were essentially un-estimable, designed to serve as a constant hemodynamic baseline against which deviations to 'K' stimuli could be effectively contrasted. Two runs were collected, each lasting 7:21 min.

During the Monetary Incentive Delay (MID) fMRI task, participants see cues that they may win or lose money ($0, $1, $5), then wait for a variable anticipatory delay period, and

finally respond to a rapidly presented target with a single button press to try to either win or avoid losing money. The interval between box offset and outcome information varied randomly from 3000-5000 msec. On incentive trials, sufficiently fast responses result in rewards, slow responses non-reward. Our extensively published version (25-30) includes three periods that represent a modification of Knutson et al.'s original fMRI MID task (31) in order to assess an extended series of information processing stages: 1) The initial presentation of the cue signaling the availability of reward; 2) the effortful response to the task and anticipatory period prior to reward notification, and 3) outcome feedback delivery. The first two phases were the conditions of greatest theoretical interest in this project as they assess conceptually different aspects of reward anticipation. The former involves the brain's response to "reward prospect" while the latter measures a combination of motivated response and waiting to learn whether that effort was sufficient to obtain reward. Both of these are relevant to different types of anticipatory information processing. Outcome delivery was not examined in this study. While it is highly germane to reward neural system function, it is less relevant to neurocognitive models of motivational impairment in ADHD that more commonly implicate how ADHD-diagnosed patients process the availability of reward and how they act in motivated contexts (32). In a pre-scan practice session, each participant's reaction times to the high-value win category ($5) were averaged and used to individually-tailor the visual cue display time so that most participants win ~⅔ of the time. To ensure optimal incentive, participants are aware they get to keep the money they earn on one run of the task determined randomly (typically in the range of ~$20-$50). Two runs were collected, each lasting 12:09 min.

Both fMRI tasks were implemented using E-Prime 2.0 (Psychology Software Tools, Inc.). The experimental stimuli were projected to the participant via a screen visible to

participants in the MRI by rear-facing mirror attached to the head coil. Behavioral responses were collected using a fiber-optic MRI-compatible response device (Current Designs, Inc.) and recorded by E-Prime for off-line analysis.

MRI Data Processing and Quality Control

Each fMRI timeseries was realigned to the mid-series volume (33), corrected for slice-timing acquisition differences (34) and spatial distortions due to inhomogeneity removed using fieldmap-based unwarping (35). Signal spikes were removed using AFNI 3dDespike (36) and volumes were automatically reoriented to stereotactic space using 3-parameter rigid body realignment. An example fMRI volume was co-registered to the MPRAGE high-resolution brain structure scan, then spatial normalization parameters mapping the T1 to MNI atlas space were applied to each fMRI volume. Each image of the resulting timeseries was written at 3 mm^3 isotropic voxel resolution then spatially smoothed with a 6 mm FWHM Gaussian kernel. Timeseries variance related to sporadic head motion was removed and the data linearly detrended using fMRIB's ICA-based Xnoiseifier (FIX v1.602 beta) (37, 38), using a training dataset constructed using a representative sample of fMRI datasets from this project with and without head motion (FIX classification threshold=20).

After omitting 10 datasets collected from left-handed ADHD participants to avoid interpretive problems related to varying hemispheric laterality, the other n=74 available ADHD fMRI datasets were matched with n=74 datasets from the non-ADHD group that had statistically equivalent mean age and sex proportion. We visually inspected the fMRI data to identify any sudden head displacements greater than a voxel length that might not be adequately controlled by ICA-FIX or during fMRI activation estimation. In the non-ADHD group, we identified 5 GNG

and 1 MID task that were omitted from further analysis due to at least one or more large displacements in both task runs. All ADHD GNG datasets were retained, but 2 ADHD MID datasets were discarded. To maximize available data, participants who had at least one good run from each task were retained (non-ADHD 1 GNG and 1 MID; ADHD 4 GNG and 6 MID). In addition, several more datasets were omitted for technical reasons (e.g., problems collecting behavioral responses, a missing MPRAGE scan that prevented T1-guided spatial normalization, etc.). The final fMRI analyses were performed using *n*=63 non-ADHD and *n*=62 ADHD GNG datasets and *n*=69 non-ADHD and *n*=62 ADHD MID datasets. For GNG, the final ADHD subgroup sizes were ADHD-EF *n*=22, ADHD-EF/REW *n*=16, ADHD-NONE *n*=24. For MID, the final ADHD subgroup sizes were ADHD-EF *n*=22, ADHD-EF/REW *n*=15, ADHD-NONE *n*=25. These final study groups did not differ in the absolute value of head displacement in the x, y, or z direction following realignment, nor did the root mean square intensity difference of successive volumes (i.e., DVARS (39)) statistically differ among groups for either fMRI task. The final fMRI subsample study groups also did not statistically differ by mean age or sex proportion.

fMRI Brain Activation Modeling and Subgroup Comparison

Individual participant activation to GNG false alarm and correctly rejected 'K' stimuli and to the two MID reward anticipation-related task phases was done in SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/). For both tasks, stimuli onsets for each experimental condition were convolved with a conventional hemodynamic response model. To segregate any remaining variance in the timeseries due to head motion that was not removed by ICA-FIX, timepoints indicated by DVARS with outlier scan-to-scan movement (75^{th} %ile + 1.5

times the inter-quartile range) were each modeled with their own regressor. All activation general linear models included a temporal derivative term, 128 sec high-pass filter, and global AR(1) model to correct for timeseries serial autocorrelation.

The fMRI group analyses had two distinct objectives. First, we needed to identify brain regions that were "abnormal" in each of the ADHD subgroups relative to non-ADHD controls. To avoid statistical disadvantages conferred by unequal cell sizes in an ANOVA, each ADHD subgroup was contrasted to non-ADHD using an SPM12 two-sample *t* test model. Multiple comparisons control for searching the entire brain volume were implemented using a clusterwise inference framework (40) requiring both a cluster-determining threshold (CDT) of $p<.005$ and extent thresholds to be surpassed for results to be considered significant. To address recent demonstrations this inference method can inflate false positive rates (41), we used updated AFNI code that corrected for inaccurate search volume space 'edge effects' and a new method to estimate noise smoothness with a non-Gaussian spatial autocorrelation function (ACF). We used mean ACF from all tested subjects' timeseries residuals as input into 3dClustSim. This approach recently has been shown to yield accurate false positive error-rate control with $p<.005$ CDT for event-related fMRI paradigms and similar FWHM smoothing values as in this study (42). Moreover, our choice of $p<.005$ cluster-determining threshold follows recent recommendations (43) to strike an acceptable balance between the need for a strong inference framework versus available statistical power in our modest-sized fMRI subsamples, the novel nature of the questions being addressed in this study, and spatial precision considerations.

Second, we wished to interrogate all these regions-of-interest (ROI) for evidence that each abnormality was specific to the subgroup in which it was found. These distinct study questions were best done using independent statistical tests to avoid interpretive ambiguities

(44). For each ROI identified by the primary fMRI analyses, we extracted average contrast values within a 3 mm radius sphere centered at the regional peak effect reported in Tables 3-5. Using these data, we conducted a secondary ANOVA to simultaneously evaluate pairwise differences among all three ADHD subgroups for the most efficient control of Type I error. We used False Discovery Rate (45) applied to all ROIs within each fMRI contrast (e.g., all ROIs within GNG response inhibition). A *post hoc* power analysis for the ADHD subgroup sample sizes indicated that this ANOVA could only detect "large" effect size differences (i.e., $f > 0.52$). As such, caution should be taken in interpreting omnibus *F* significance levels, particularly when uncorrected group differences fell under the conventional $p<.05$ cutoff but failed to be confirmed after correcting for multiple comparisons. Finally, to complement this analysis we examined the question of ADHD subgroup-specificity by conducting a conjunction analysis, following well-described methodology (25). For each fMRI task condition examined, three contrasts were specified to test for simple study group differences between non-ADHD and one of the three ADHD subgroups. In the approach taken in SPM12, the conjunction corresponds to a significant sum of all the individual contrast effects, if and only if there are no significant differences among them. These estimates are represented as a statistical map and thresholded at $p <.05$ uncorrected. A liberal threshold was deliberately chosen to assess whether any evidence could be found for abnormalities shared by all three subgroups, even if such evidence would not hold up to conventional corrections for multiple comparisons.

fMRI Task Behavioral Performance

For the GNG task, a series of *t* tests found that the number of false alarm errors in non-ADHD was lower than found in both the ADHD-EF ($t = -2.502$, $p = .014$) and ADHD-EF/REW

(t = -3.686, p <.001), but not the ADHD-NONE. However, secondary ANOVA among the 3 ADHD subgroups indicated this false alarm performance difference was non-significant ($F_{2,59}$ = 1.975, p = .148). The number of misses to 'X' stimuli did not differ between non-ADHD or any of the ADHD subgroups.

The results of t test analyses for MID task performance data showed that none of the ADHD subgroups differed from non-ADHD in average reaction time to incentivized Win trials. Only the ADHD-EF/REW differed from non-ADHD average monetary winnings per run (t = 1.749, p = .010), where ADHD-EF/REW won less money ($10.29 vs. $12.96). However, ANOVA failed to detect differences among the ADHD subgroups for average winnings.

Table S1. Neuropsychological test performance differences between ADHD-diagnosed adolescents and non-ADHD control participants as evaluated by two-sample *t* test.

	Controls *n*=134	ADHD *n*=117	*p*
Delay Discounting (AUC)	0.382 (0.28)	0.325 (0.22)	.070 †
Experiential Discounting (AUC)	0.639 (0.17)	0.602 (0.15)	.072 †
SKIP average IRT (*sec*)	13.48 (21.26)	7.451 (15.94)	.011
Stop Signal RT (*msec*)	277.936 (60.35)	306.901 (77.76)	.001
IMT Impulsivity (ratio)	0.506 (0.19)	0.652 (0.2)	<.001
DMT Impulsivity (ratio)	0.506 (0.24)	0.761 (0.42)	<.001
MMF Errors (#)	17.101 (10.98)	30.08 (16.03)	<.001
CPT-II Commissions (#)	20.145 (6.61)	22.693 (7.13)	.004

† statistical trend $p < .10$.

AUC – Area under the curve

Table S2. Factor structure estimated separately for ADHD and non-ADHD study groups using principal component analysis. Component loadings >.300 are listed.

	Non-ADHD			ADHD		
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
Delay Discounting (AUC)	-	-0.695	-	-	-	0.765
Experiential Discounting (AUC)	-	-	0.738	-0.363	-0.423	0.566
SKIP average IRT (*sec*)	-	0.757	-	-	0.577	-
Stop Signal RT (*msec*)	0.549	0.392	-	-	0.606	0.492
IMT Impulsivity (ratio)	0.731	-	-	0.815	-	-
DMT Impulsivity (ratio)	0.632	-	0.499	-	0.727	-
MMF Errors (#)	-	-	0.537	-	-	-
CPT-II Commissions (#)	0.712	-	-	0.798	-	-

AUC – Area under the curve

Table S3. Taxometric-derived ADHD subgroup characteristics from other available neuropsychological test scores compared to non-ADHD control in a series of two-sample *t* tests and results from one-way Analysis of Variance (ANOVA) directly comparing the ADHD subgroups.

	Controls *n*=134	ADHD-EF *n*=40	ADHD-EF/REW *n*=31	ADHD-NONE *n*=46	ADHD Subgroup ANOVA *p*
CPT-II Omissions	6.331 (1.55)	18.141 (2.78) *	22.911 (3.18) *	12.528 (2.61) *	*ns*
CPT-II Hit Reaction Time SE	6.61 (0.44)	11.731 (0.79) *	11.859 (0.9) *	8.386 (0.74) *	.019
CPT-II Variability	12.439 (1.71)	29.332 (3.08) *	27.495 (3.51) *	19.121 (2.88) *	*ns*
CPT-II Hit Reaction Time	336.167 (4.73)	369.507 (8.5) *	371.178 (9.7) *	375.429 (7.96) *	*ns*
IMT Latency	454.43 (5.04)	439.82 (9.07)	432.12 (10.35)	493.06 (8.5) *	<.001
DMT Latency	483.67 (7.46)	445.83 (13.42) *	424.94 (15.32) *	514.44 (12.58) *	<.001
IMT d’	1.288 (0.05)	0.777 (0.09) *	0.508 (0.1) *	1.086 (0.08) *	<.001
DMT d’	1.545 (0.07)	0.743 (0.12) *	0.427 (0.14) *	1.191 (0.11)	<.001
CPT-II d’	0.424 (0.03)	0.287 (0.06) *	0.164 (0.06) *	0.546 (0.05) *	<.001
IMT beta	0.653 (0.03)	0.752 (0.05)	0.806 (0.06) *	0.816 (0.05)	*ns*
DMT beta	0.623 (0.08)	0.81 (0.14)	1.262 (0.17) *	0.785 (0.14)	*ns*
CPT-II beta	0.554 (0.04)	0.627 (0.08)	0.77 (0.09) *	0.688 (0.07)	*ns*
Probabilistic Discounting	0.396 (0.02)	0.377 (0.03)	0.402 (0.03)	0.395 (0.03)	*ns*

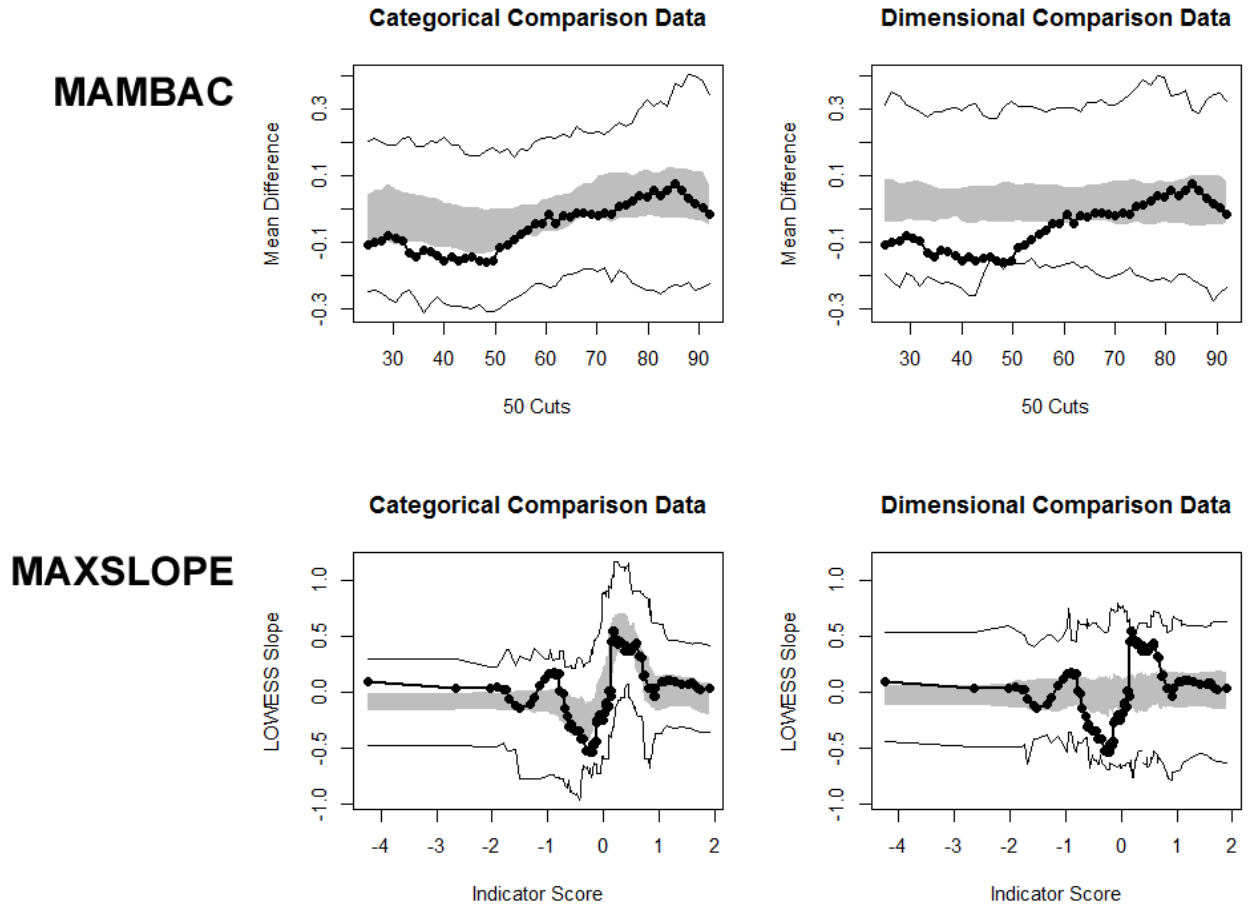

Figure S1. Comparison of simulated dimensional and categorical comparison data fit to the taxometric model determined in this study using MAMBAC and MAXSLOPE methods for consistency checks.

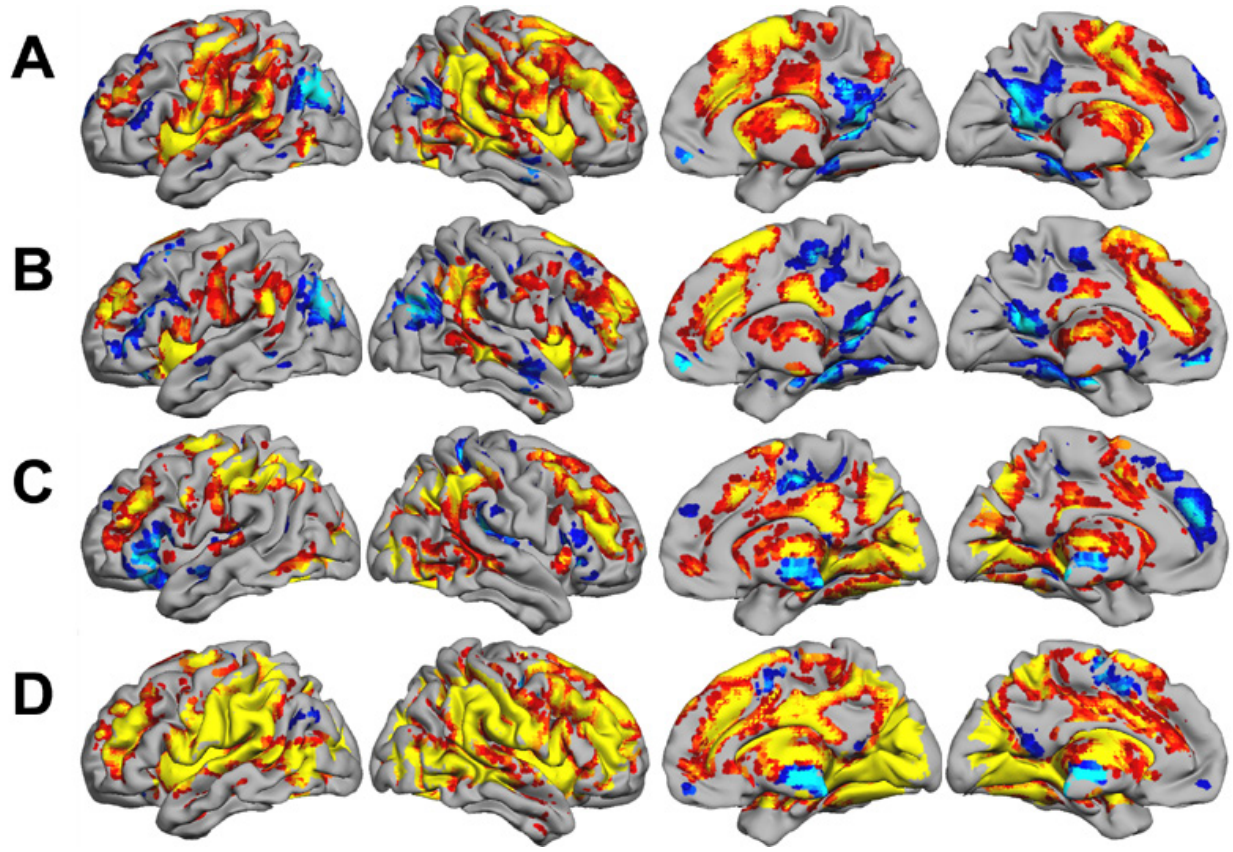

Figure S2. This figure depicts simple activation profiles ($p<.01$ uncorrected) across all ADHD and non-ADHD participants who underwent fMRI for the GNG and MID tasks as whole brain renderings on a representative anatomical atlas. Red-yellow indicates brain regions with positive hemodynamic signal change relative to an unmodeled implicit baseline; blue-light blue represents negative change. A) Brain activity elicited when NoGo 'K' stimuli were correctly inhibited. B) Brain activity elicited during "false alarm" errors to NoGo 'K' stimuli when participants incorrectly made a motor response. C) Brain activity elicited when participants were presented a cue indicating they might win $5 or $1. D) Brain activity elicited when participants made a speeded button press and awaited the outcome of their effort.

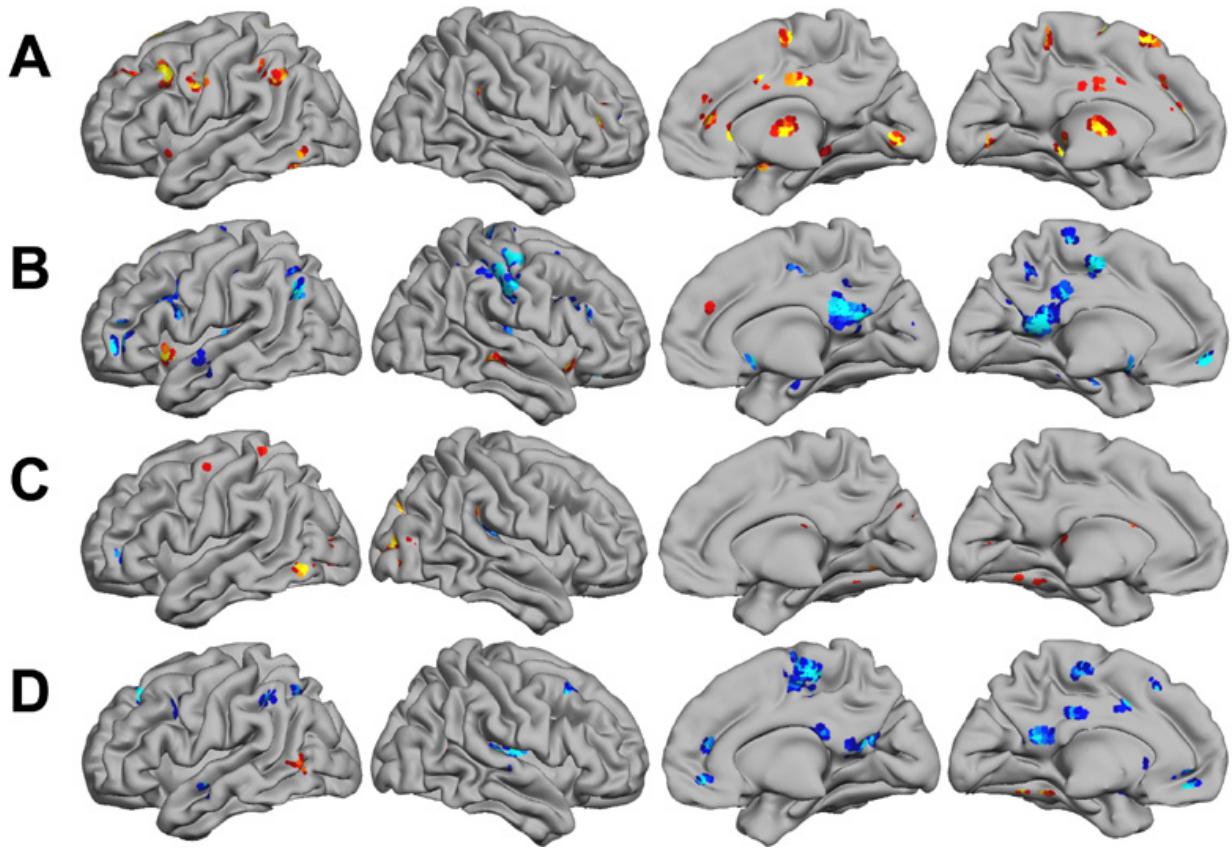

Figure S3. This figure depicts differences (p<.01 uncorrected) between non-ADHD and ADHD-diagnosed participants on the four conditions of interest on the GNG and MID fMRI tasks (A-D). All regions depicted survive a "whole brain" correction for multiple comparisons from Monte Carlo simulation. Red-yellow indicates brain regions where non-ADHD had greater hemodynamic response than ADHD; blue-light blue represents greater activation in ADHD compared to non-ADHD. A) Correctly-inhibited NoGo 'K' stimuli. B) False alarm responses to NoGo 'K' stimuli. C) Brain activity to cue indicating the availability of monetary reward. D) Brain activity when participants exerted effort to obtain reward and awaited outcome notification.

Supplemental Information References

1. Brown TE (2001) Brown Attention-Deficit Disorder Scales for Adolescents and Adults. (Harcourt Assessment, San Antonio, TX).

2. Achenbach TM & Edelbrock C (1983) *Manual of the Child Behavior Checklist and Revised Child Behavior Profile* (University of Vermont, Department of Psychiatry, Burlington).

3. Beck AT, Steer RA, & Brown GK (1996) *Beck Depression Inventory—II (BDI–II)* (Harcourt Assessment, Inc., San Antonio, TX).

4. March JS (*Multidimensional Anxiety Scale for Children (2nd Edition).*

5. Fagerstrom KO & Schneider NG (1989) Measuring nicotine dependence: a review of the Fagerstrom Tolerance Questionnaire. *J Behav Med* 12(2):159-182.

6. Moberg DP (2005) Screening for Alcohol and Other Drug Problems using the Adolescent Alcohol and Drug Involvement Scale (AADIS). (University of Wisconsin Medical School, Center for Health Policy and Program Evaluation).

7. Weissman MM, *et al.* (2000) Brief screening for family psychiatric history: the family history screen. *Arch Gen Psychiatry* 57(7):675-682.

8. Alderson RM, Rapport MD, & Kofler MJ (2007) Attention-deficit/hyperactivity disorder and behavioral inhibition: a meta-analytic review of the stop-signal paradigm. *J Abnorm Child Psychol* 35(5):745-758.

9. Doyle AE (2006) Executive functions in attention-deficit/hyperactivity disorder. *J Clin Psychiatry* 67 Suppl 8:21-26.

10. Patros CH, *et al.* (2016) Choice-impulsivity in children and adolescents with attention-deficit/hyperactivity disorder (ADHD): A meta-analytic review. *Clin Psychol Rev* 43:162-174.

11. Conners CK, Epstein JN, Angold A, & Klaric J (2003) Continuous performance test performance in a normative epidemiological sample. *J Abnorm Child Psychol* 31(5):555-562.

12. Logan GD, Cowan WB, & Davis KA (1984) On the ability to inhibit simple and choice reaction time responses: a model and a method. *J Exp Psychol Hum Percept Perform* 10(2):276-291.

13. Verbruggen F, Logan GD, & Stevens MA (2008) STOP-IT: Windows executable software for the stop-signal paradigm. *Behav Res Methods* 40(2):479-483.

14. Logan GD, Van Zandt T, Verbruggen F, & Wagenmakers EJ (2014) On the ability to inhibit thought and action: general and special theories of an act of control. *Psychological review* 121(1):66-95.

15. Dougherty DM, Marsh DM, & Mathias CW (2002) Immediate and delayed memory tasks: a computerized behavioral measure of memory, attention, and impulsivity. *Behav Res Methods Instrum Comput* 34(3):391-398.

16. Kagan J, Rosman BL, Day D, Albert J, & Phillips W (1964) Information processing in the child: signifiance of analytic and reflective attitudes. *Psychological Monographs* 78:1-37.

17. Kirby KN (2009) One-year temporal stability of delay-discount rates. *Psychon Bull Rev* 16(3):457-462.

18. Myerson J, Green L, & Warusawitharana M (2001) Area under the curve as a measure of discounting. *J Exp Anal Behav* 76(2):235-243.

19. Reynolds B & Schiffbauer R (2004) Measuring state changes in human delay discounting: an experiential discounting task. *Behav Processes* 67(3):343-356.

20. Smits RR, Stein JS, Johnson PS, Odum AL, & Madden GJ (2013) Test-retest reliability and construct validity of the Experiential Discounting Task. *Exp Clin Psychopharmacol* 21(2):155-163.

21. Dougherty DM, Mathias CW, Marsh DM, & Jagar AA (2005) Laboratory behavioral measures of impulsivity. *Behav Res Methods* 37(1):82-90.

22. Meehl PE & Yonce LJ (1994) Taxometric analysis: I. Detecting taxonicity with two wauntifative indicators using means above and below a sliding cut (MAMBAC procedure). *Psychological Reports* 74:1059-1274.

23. Grove WM (2004) The MAXSLOPE taxometric procedure: Mathematical derivation, parameter estimation, consistency tests. *Psychological Reports* 95:517-515.

24. Ruscio J & Ruscio AM (2004) A conceptual and methodological checklist for conducting a taxometric investigation. *Behavior Therapy* 35:403-447.

25. Balodis IM, *et al.* (2014) A pilot study linking reduced fronto-Striatal recruitment during reward processing to persistent bingeing following treatment for binge-eating disorder. *Int J Eat Disord* 47(4):376-384.

26. Balodis IM, *et al.* (2012) Diminished frontostriatal activity during processing of monetary rewards and losses in pathological gambling. *Biol Psychiatry* 71(8):749-757.

27. Balodis IM, *et al.* (2013) Monetary reward processing in obese individuals with and without binge eating disorder. *Biol Psychiatry* 73(9):877-886.

28. Jia Z, *et al.* (2011) An initial study of neural responses to monetary incentives as related to treatment outcome in cocaine dependence. *Biol Psychiatry* 70(6):553-560.

29. Patel KT, *et al.* (2013) Robust changes in reward circuitry during reward loss in current and former cocaine users during performance of a monetary incentive delay task. *Biol Psychiatry* 74(7):529-537.

30. Urban NB, *et al.* (2012) Imaging human reward processing with positron emission tomography and functional magnetic resonance imaging. *Psychopharmacology (Berl)* 221(1):67-77.

31. Knutson B, *et al.* (2004) Amphetamine modulates human incentive processing. *Neuron* 43(2):261-269.

32. Luman M, Tripp G, & Scheres A (2010) Identifying the neurobiology of altered reinforcement sensitivity in ADHD: a review and research agenda. *Neurosci Biobehav Rev* 34(5):744-754.

33. Jenkinson M, Bannister P, Brady M, & Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17(2):825-841.

34. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, & Smith SM (2012) Fsl. *Neuroimage* 62(2):782-790.

35. Andersson JL, Hutton C, Ashburner J, Turner R, & Friston K (2001) Modeling geometric deformations in EPI time series. *Neuroimage* 13(5):903-919.

36. Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29(3):162-173.

37. Griffanti L, *et al.* (2014) ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage* 95:232-247.

38. Salimi-Khorshidi G, *et al.* (2014) Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* 90:449-468.

39. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, & Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59(3):2142-2154.

40. Forman SD, *et al.* (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med* 33(5):636-647.

41. Eklund A, Nichols TE, & Knutsson H (2016) Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci U S A* 113(28):7900-7905.

42. Cox RW, Chen G, Glen DR, Reynolds RC, & Taylor PA (in press) FMRI clustering in AFNI: False positive rates redux. *Brain Connectivitiy*.

43. Carter CS, Lesh TA, & Barch DM (2016) Thresholds, power, and sample sizes in clinical neuroimaging. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 1(2):99-100.

44. Kriegeskorte N, Simmons WK, Bellgowan PS, & Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nature neuroscience* 12(5):535-540.

45. Benjamini Y & Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. Roy Stat Soc, Ser B* 57:289-300.