Washington University School of Medicine
Digital Commons@Becker

Open Access Publications

2013

Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data

Nico U.F. Dosenbach
Washington University School of Medicine in St. Louis

Bradley L. Schlaggar
Washington University School of Medicine in St. Louis

et al

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation
Dosenbach, Nico U.F.; Schlaggar, Bradley L.; and et al, ,"Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data." Frontiers in Systems Neuroscience.6,. 80. (2013).
http://digitalcommons.wustl.edu/open_access_pubs/2234

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.

frontiers in
SYSTEMS NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 04 February 2013
doi: 10.3389/fnsys.2012.00080

Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data

Damien A. Fair[1,2]*, Joel T. Nigg[1,2], Swathi Iyer[1,2], Deepti Bathula[1,2,3], Kathryn L. Mills[1,2], Nico U. F. Dosenbach[4], Bradley L. Schlaggar[4], Maarten Mennes[5], David Gutman[5], Saroja Bangaru[5], Jan K. Buitelaar[6], Daniel P. Dickstein[7], Adriana Di Martino[5], David N. Kennedy[8], Clare Kelly[5], Beatriz Luna[9], Julie B. Schweitzer[10], Katerina Velanova[9], Yu-Feng Wang[11,12], Stewart Mostofsky[13,14], F. Xavier Castellanos[5,15] and Michael P. Milham[15,16]*

[1] *Department of Behavioral Neuroscience, Advanced Imaging Research Center, Oregon Health and Science University, Portland, OR, USA*
[2] *Department of Psychiatry, Advanced Imaging Research Center, Oregon Health and Science University, Portland, OR, USA*
[3] *Department of Computer Science and Engineering, Indian Institute of Technology Ropar, Rupnagar, Punjab, India*
[4] *Department of Neurology, Washington University, St. Louis, MO, USA*
[5] *Phyllis Green and Randolph Cowen Institute for Pediatric Neuroscience, NYU Langone Medical Center, New York, NY, USA*
[6] *Department of Cognitive Neuroscience, Radboud University Nijmegen Medical Center, Nijmegen, Netherlands*
[7] *PediMIND Program, Bradley Hospital, Brown University School of Medicine, East Providence RI, USA*
[8] *University of Massachusetts Medical Center, Worcester, MA, USA*
[9] *Department of Psychiatry, University of Pittsburgh, PA, USA*
[10] *Departments of Psychiatry, Behavioral Sciences, and the MIND Institute, University of California Davis School of Medicine, Sacramento, CA, USA*
[11] *Institute of Mental Health, Peking University, Beijing, China*
[12] *Key Laboratory of Mental Health, Ministry of Health, Peking University, Beijing, China*
[13] *Kennedy Krieger Institute, Baltimore, MD, USA*
[14] *Johns Hopkins University, Baltimore, MD, USA*
[15] *Nathan Kline Institute, Orangeburg, NY, USA*
[16] *Center for the Developing Brain, Child Mind Institute, New York, NY, USA*

Edited by:
Ranulfo Romo, Universidad Nacional Autónoma de México, Mexico

Reviewed by:
Valentin Dragoi, University of Texas Medical School at Houston, USA
David C. Somers, Boston University, USA

Correspondence:
Damien A. Fair, Department of Psychiatry, Oregon Health and Science University, 3181 SW Sam Jackson Park Road UHN88, Portland, OR 97239, USA.
e-mail: faird@ohsu.edu

Michael P. Milham, Child Mind Institute, Center for the Developing Brain, 445 Park Avenue, New York, NY 10022, USA.
e-mail: michael.milham@childmind.org

In recent years, there has been growing enthusiasm that functional magnetic resonance imaging (MRI) could achieve clinical utility for a broad range of neuropsychiatric disorders. However, several barriers remain. For example, the acquisition of large-scale datasets capable of clarifying the marked heterogeneity that exists in psychiatric illnesses will need to be realized. In addition, there continues to be a need for the development of image processing and analysis methods capable of separating signal from artifact. As a prototypical hyperkinetic disorder, and movement-related artifact being a significant confound in functional imaging studies, ADHD offers a unique challenge. As part of the ADHD-200 Global Competition and this special edition of Frontiers, the ADHD-200 Consortium demonstrates the utility of an aggregate dataset pooled across five institutions in addressing these challenges. The work aimed to (1) examine the impact of emerging techniques for controlling for "micro-movements," and (2) provide novel insights into the neural correlates of ADHD subtypes. Using support vector machine (SVM)-based multivariate pattern analysis (MVPA) we show that functional connectivity patterns in individuals are capable of differentiating the two most prominent ADHD subtypes. The application of graph-theory revealed that the Combined (ADHD-C) and Inattentive (ADHD-I) subtypes demonstrated some overlapping (particularly sensorimotor systems), but unique patterns of atypical connectivity. For ADHD-C, atypical connectivity was prominent in midline default network components, as well as insular cortex; in contrast, the ADHD-I group exhibited atypical patterns within the dlPFC regions and cerebellum. Systematic motion-related artifact was noted, and highlighted the need for stringent motion correction. Findings reported were robust to the specific motion correction strategy employed. These data suggest that resting-state functional connectivity MRI (rs-fcMRI) data can be used to characterize individual patients with ADHD and to identify neural distinctions underlying the clinical heterogeneity of ADHD.

Keywords: ADHD, functional connectivity, support vector machines, RDoC, research domain criteria

INTRODUCTION

Brain imaging has increasingly become a useful tool in modern medicine. Most notably, magnetic resonance imaging (MRI) has emerged as an accurate and reliable approach to identifying abnormalities characteristic of congenital, neoplastic, ischemic, inflammatory, metabolic, and infectious processes in the brain. Unfortunately, the clinical utility of imaging is markedly diminished when considering conditions that are not accompanied by gross structural or inflammatory abnormalities. Chronic pain syndromes, movement disorders, and, in particular, psychiatric illnesses, have thus far remained without scientifically founded clinical benefit from the introduction of brain imaging into medical practice (Matthews et al., 2006). After two decades of development, there is now growing enthusiasm that functional MRI could achieve clinical utility for a broad range of neuropsychiatric disorders.

This growing enthusiasm stems, in part, from the emergence of resting-state functional connectivity MRI (rs-fcMRI). Originally described by Biswal et al. (1995), rs-fcMRI is based on the discovery that spontaneous neural activity (Biswal et al., 1995; Leopold et al., 2003; Nir et al., 2006; Scholvinck et al., 2010) is associated with correlated low-frequency ($<\sim$0.1 Hz) blood oxygen level-dependent (BOLD) signal fluctuations in functionally related brain regions at rest (Biswal et al., 1995). By cross-correlating the time series of a particular brain region (seed region) with other regions or voxels, one can determine which regions are "functionally connected." Importantly, rs-fcMRI can be used during sleep as well as during sedation (Fukunaga et al., 2006, 2008; Vincent et al., 2007; Greicius, 2008; Horovitz et al., 2008); it yields consistent results across subjects, scans, and days (van de Ven et al., 2004; Damoiseaux et al., 2006; Shehzad et al., 2009; Van Dijk et al., 2010), and rs-fcMRI results are remarkably reliable across imaging centers (Biswal et al., 2010). These features make rs-fcMRI an attractive measure for translational and clinical applications. As highlighted in the ADHD-200 competition (ADHD-200-Consortium, 2012), and this special edition highlighting the work of the competitors, the race for these sorts of applications is currently underway (ADHD-200-Consortium, 2012).

Despite the enthusiasm, two key rate-limiting steps remain for the advancement of functional neuroimaging approaches in the clinical realm. First is the acquisition of large-scale datasets capable of clarifying the marked heterogeneity that exists in psychiatric disorders—even within a single diagnostic category. The diagnosis of Attention Deficit Hyperactivity disorder (ADHD) provides a salient example of clinical heterogeneity, as DSM-IV distinguishes between three distinct subtypes, including predominantly hyperactive/impulsive (relatively infrequent), predominantly inattentive, and combined (most frequent among children). While the behavioral literature has long struggled with the challenges of identifying commonalities and differences among the subtypes, imaging studies have generally ignored these distinctions. The second challenge is the development of image processing and analytical methods capable of separating signal from artifact, enabling both the characterization of pathologic processes underlying a given disorder, and detection of their presence in an individual. As a prototypical hyperkinetic disorder, ADHD presents exceptional challenges for fMRI-based research, due to importance of addressing the increased prevalence of motion in this population, which can artificially produce or obscure ADHD-related differences in rs-fcMRI metrics (Power et al., 2012a; Satterthwaite et al., 2012; Van Dijk et al., 2012).

In response to these challenges, here we use an aggregate dataset pooled across five institutions [resulting in quality-controlled rs-fcMRI scans for 455 typically developing children (TDC) and 193 children with ADHD] to provide novel insights into the neural correlates of clinical heterogeneity in ADHD. Specifically, we use support vector machine (SVM)-based multivariate pattern analysis (MVPA) to identify those functional connections in the brain that, collectively, are capable of differentiating the DSM-IV Inattentive and Combined ADHD subtypes (termed presentations in DSM-5) from one another, as well as from typically developing controls (we omit the hyperactive-impulsive subtype due to its relative rarity in the age ranges studied). We believe that such pattern analytic approaches may prove to be advantageous in the examination of ADHD, given the growing consensus that the neural correlates of ADHD are distributed in nature, rather than being explained by abnormalities in any specific connection or region.

In this work we begin with a methodological aim to assess several techniques aimed at controlling for micro-movements in a TDC sample. We focus this analysis on results related to short and long-range functional connectivity, as recent works suggest that findings related to these types of connections can be augmented by micro-movements (Power et al., 2012a; Satterthwaite et al., 2012; Van Dijk et al., 2012). Using the strongest methods from the first aim, we then follow-up this examination to investigate subtype heterogeneity and the predictive capacity of rs-fcMRI in a large group of TDC vs. a large group of children with ADHD.

METHODS

PARTICIPANTS AND MEASURES

Data from Brown University, Beijing Normal University, Kennedy Krieger Institute, New York University Child Study Center, Washington University at St. Louis, and Oregon Health and Science University were aggregated for youth ages 7–14 years. The resulting dataset comprised 455 typically developing control subjects and 193 subjects with a DSM-IV-TR diagnosis of ADHD. A summary of the demographic characteristics for each sample is provided in **Table 1**. Informed written consent and assent were obtained for all participants, and procedures complied with the Human Investigation Review Board at respective universities. As data for this investigation were aggregated retrospectively (as a large collaborative effort), slightly different ADHD assessment protocols were used across institutions. These procedures are detailed in Appendix text.

DATA ACQUISITION

All participants were scanned on 3.0 Tesla scanners using standard resting-connectivity T2*-weighted echo-planar imaging (details for each institution are provided in Appendix text). All imaging data used is publicly available at the Neuroimaging

Table 1 | Sample characteristics.

	N	Age (mean)	% Female	FD mean	ADHD-I				ADHD-C			
					N	Age (mean)	% Female	FD mean	*N*	Age (mean)	% Female	FD mean
Total	455	14.39	50	0.10	80	11.45	27	0.11	112	10.31	19	0.11
Brown	6	12.38	83	0.11	1	14.70	0	0.00	5	11.86	80	0.11
NYU	65	15.54	54	0.10	15	10.54	47	0.12	37	9.95	16	0.11
Beijing Normal	176	13.63	44	0.10	48	12.14	17	0.11	40	11.25	05	0.11
JHU	70	13.78	53	0.09	5	11.27	20	0.08	14	9.77	50	0.09
OHSU	56	13.94	59	0.11	11	9.15	45	0.11	16	8.80	13	0.11
WashU	82	16.09	48	0.11	–	–	–	–	–	–	–	–

FD mean reflects the mean frame displacement of the remaining frames after censoring (see Methods).

Informatics Tools and Resources Clearinghouse (NITRC), see http://fcon_1000.projects.nitrc.org/indi/adhd200.

PREPROCESSING

All functional images were preprocessed in the same manner to reduce artifacts (Miezin et al., 2000). These steps included: (1) removal of a central spike caused by MR signal offset, (2) correction of odd vs. even slice intensity differences attributable to interleaved acquisition without gaps, (3) correction for head movement within and across runs [also see Power et al. (2012a)], and (4) within-run intensity normalization to a whole brain mode value of 1000. Atlas transformation of the functional data was computed for each individual via the MPRAGE scan. Each run then was resampled in atlas space (Talairach and Tournoux, 1988) on an isotropic 3 mm grid, combining movement correction and atlas transformation in one interpolation (Lancaster et al., 1995). All subsequent operations were performed on the atlas-transformed volumetric time series.

Functional connectivity preprocessing followed prior methods (Fox et al., 2005; Fair et al., 2007a,b, 2008, 2009). These steps included: (1) a temporal band-pass filter ($0.009\,\text{Hz} < f < 0.08\,\text{Hz}$) and spatial smoothing (6 mm full width at half maximum), (2) regression of the whole brain signal averaged over the whole brain, (3) regression of ventricular signal averaged from ventricular region of interest (ROI), and (4) regression of white matter signal averaged from white matter ROI. Regression of first order derivative terms for the whole brain, ventricular, and white matter signals were also included in the correlation preprocessing. These preprocessing steps are thought to reduce spurious variance unlikely to reflect neuronal activity (Fox and Raichle, 2007).

TRADITIONAL MOTION PARAMETERS AND CORRECTION

In a typical functional connectivity experiment, motion is addressed by excluding participants with high levels of movement (using various criteria), and then removing movement-related signal via a linear regression of preprocessed data on the 6 motion parameters (i.e., rotation and translation) for remaining participants. In some instances, samples are matched for movement [via parameters such as root mean square (RMS)] (Fair et al., 2007a; Dosenbach et al., 2010). However, these approaches involve potentially problematic assumptions. The first is that the traditional calculations of the 6 motion parameters (which are typically generated relative to a within-run reference frame) are tightly related to abrupt motion-related changes in the BOLD signal. The second is that there is a linear relationship between changes in the BOLD signal and abrupt motion in the scanner. Three recent reports (Power et al., 2012a; Satterthwaite et al., 2012; Van Dijk et al., 2012) suggest that these assumptions are likely incorrect and that traditional motion correction does not adequately control for the changes in signal intensity that accompany abrupt changes in head position.

With this concern in mind, we attempted several motion correction procedures (described below). At the first level of correction (i.e., traditional motion correction), motion was measured relative to a reference frame (in this case, the middle frame of a BOLD run) and quantified using an analysis of head position based on rigid body translation and rotation. This procedure results in the rigid body transform defined by 6 motion parameters (3 translation, 3 rotation) typically generated by most functional MRI software tools. Traditional motion correction procedures in fMRI-based functional connectivity studies, as well as in many task-based fMRI studies, use these 6 parameters as regressors in preprocessing to remove potential motion-related artifact. This step was included in most analyses below.

In addition, in an effort to remove participants with egregious motion, we began our analysis by filtering those subjects with high movement runs based on RMS. The data derived from the 6 motion parameters needed to realign head movement on a frame-by-frame basis were calculated as RMS values for translation and rotation in the x, y, and z planes in millimeters. Total RMS values were calculated on a run-by-run basis for each participant. Participant's BOLD runs with movement exceeding 1.5 mm RMS were removed.

FRAME-TO-FRAME (VOLUME-TO-VOLUME) MOTION PARAMETERS

In an effort to examine motion from volume-to-volume, two additional motion parameters were examined. The first, based on framewise displacement (FD), was first introduced by Power et al. (2012a). This variable measures movement of any given frame relative to the previous frame, as opposed to relative to the reference frame (as above). Thus, the method yields a 6 dimensional time series representing frame-to-frame motion, as

described by $FD_i = |\Delta d_{ix}| + |\Delta d_{iy}| + |\Delta d_{iz}| + |\Delta \alpha_i| + |\Delta \beta_i| + |\Delta \gamma_i|$, where $\Delta d_{ix} = d_{(i-1)x} - d_{ix}$, and similarly for the other five rigid body parameters $[d_{ix}\, d_{iy}\, d_{iz}\, \alpha_i\, \beta_i\, \gamma_i]$. In essence, this formula sums the absolute values of volume-by-volume changes in the six rigid body parameters (Note: Rotational displacements for this method are first converted from degrees to millimeters by calculating surface displacement on a sphere of radius 50 mm, the approximate distance from the cerebral cortex to the center of the head).

A second measure likely to reflect direct BOLD-related deviations secondary to movement is termed DVARS (the RMS of the derivatives of the differentiated timecourses of every brain voxel for each acquired volume), originally described by Smyser et al. (2010). DVARS quantifies volume-to-volume BOLD signal change, thus capturing large deviations produced by phenomena that impact the brain on a global scale—head motion being a major contributor to fluctuations in DVARS. This measure is based on the fact that abrupt head displacement typically manifests as signal loss in echo-planar imaging (Smyser et al., 2010). Thus, a logical measurement sensitive to sudden changes in head position is the whole brain signal change measured by DVARS (Power et al., 2012a). DVARS is computed by aggregating voxel-wise volume-by-volume backwards differences in the BOLD signal described by: $\text{DVARS}\ (\Delta I_i)_t = \sqrt{\left\langle [\Delta I_i\,(\vec{x})]^2 \right\rangle} = \sqrt{\left\langle [I_i\,(\vec{x}) - I_{i-1}\,(\vec{x})]^2 \right\rangle}$, where $I_i(\vec{x})$ is image intensity at locus $\vec{x}$ on frame i and angle brackets denote the spatial average over the whole brain. Because frame-to-frame changes in signal intensity related to movement are significantly greater than those caused by neurophysiologic changes in the BOLD signal, this measure provides a natural parameter with which to directly examine the relationship of movement measurements and the BOLD response (Smyser et al., 2010).

REGION OF INTEREST SELECTION

We selected 160 regions of interest (ROIs) based on prior work by Dosenbach et al. (2010) (see **Figure A7**). These regions were selected based on their use to develop maturation indices (fcMI) in a previous study (Dosenbach et al., 2010). This set of regions originated from a series of five meta-analyses, focused on error-processing, default-mode (task-induced deactivations), memory, language, and sensorimotor functions (Dosenbach et al., 2010). The functionally defined ROIs were used as a "best guess" estimate of the underlying functional area architecture across the brain (Fair and Schlaggar, 2008).

Network categorization of each ROI was based on labels designated in a previous report (Dosenbach et al., 2010). This categorization stemmed from a community detection procedure conducted on combined correlation matrices across adult subjects [e.g., see Fair et al. (2009)]. The modularity optimization algorithm of Newman was used (Newman, 2006). The modules (i.e., communities) used to categorize regions had a high quality index (Q) and were the most resistant to perturbation by randomization, measured by variation of information (VOI) (Karrer et al., 2008). For **Figure A7**, the network assignments were re-examined based on the current adult and child datasets. In this instance, the weight conserving community detection algorithm used was based on the work by Rubinov and Sporns (2011).

COMPUTATION OF SINGLE SUBJECT CORRELATION MATRICES

SVR analyses in the manuscript were conducted on single subject correlation matrices derived from the above mentioned 160 *a priori* ROIs (10 mm diameter spheres). The resting-state BOLD time series were correlated region by region for each participant across the full length of the resting time series, creating 455 square correlation matrices (160 × 160), one for each subject (Dosenbach et al., 2007; Fair et al., 2007a, 2008, 2009). For those motion correction strategies in which motion covariates were included in the calculation, the resultant matrices for each participant represent partial correlations (after accounting for motion).

AGE RELATIONSHIPS

For all statistical comparisons, *r*-values within matrices were first normalized using the Fisher r-to-z transformation. Functional connections of the brain most strongly associated with age were then determined by cross correlating each connection with subject age (in one instance partial correlations were generated using mean movement parameters as a covariate—see below). Connections were corrected with the Benjamini and Hochberg False Discovery Rate (Benjamini et al., 2001). Connection distances for those links that significantly correlated with age were calculated in terms of Euclidean distance (i.e., vector distance) between the center of mass coordinates of each region. These values were then separated for those connections that got stronger with age or weaker with age. *Post-hoc* analyses for these same connections using one of the selected movement-related procedures below (i.e., Procedure 8) was conducted for **Figure A2** (site by site comparisons).

SUPPORT VECTOR-BASED MULTIVARIATE PATTERN ANALYSIS

In the present study SVM and support vector regression (SVR)-based MVPA were used to make predictions about brain maturity and disease status in individual subjects (e.g., Norman et al., 2006; Dosenbach et al., 2010; Rizk-Jackson et al., 2011). The approach used in the present analysis was similar to that used in a previous report (Dosenbach et al., 2010).

SVM-based MVPA is a supervised classification algorithm rooted in statistical learning theory. Conceptually, input vectors are mapped to a higher dimensional feature space using special non-linear functions called kernels. Classification is performed by constructing a hyperplane in the feature space that optimally discriminates between two classes of the training data by maximizing the margin between two data clusters.

Given a training set of the form $(\mathbf{x}_i, y_i)$ where the vectors $\mathbf{x}_i$ are data points and y_i are the class labels, the SVMs require the solution to the following optimization problem:

$$\min_{\boldsymbol{w},\, b,\, \xi} \ \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_{i=i}^{n}\xi_i$$

subject to $y_i(\boldsymbol{w} \times \boldsymbol{x}_i + b) \geq 1 - \xi_i$ and $\xi_i \geq 0$

where ξ_i are the slack variables, measuring the degree of a data point's misclassification, $\boldsymbol{w}$ are the weights defining the hyperplane and $C > 0$ is the penalty parameter of the error term. The resultant decision function implemented by SVM can be written as:

$$f(x) = sign\left(\sum_{i=1}^{n} y_i \alpha_i K(\boldsymbol{x},\ \boldsymbol{x}_i) + b\right)$$

where $K(x_i,\ x_j)$ is the kernel function. In our work, we use the radial basis kernel given by:

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = exp\left(-\frac{\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|^2}{2\sigma^2}\right),$$

SVMs are inherently two-class classifiers. Multiclass SVM aims to handle the K-class pattern classification problem by reducing a single multiclass problem into multiple binary classification problems. The most common method for such reduction is to build a set of *one-vs.-rest* binary classifiers that distinguish one of the classes from the rest. Another strategy is to build a set of *one-vs.-one* classifiers that distinguish between every pair of classes. For the *one-vs.-one* approach, classification is done by max-wins voting strategy that chooses the class that is selected by the most classifiers. For the *one-vs.-rest* case (used in this work), classification of new instances is done by a winner-takes-all strategy, in which the classifier with the highest output function assigns the class. SVM classifications used a soft margin $C = 1$, and a radial basis function with $\sigma = 2$. In brief, for all SVR classifications we used epsilon-insensitive SVRs. Parameters were set with C = Infinity, epsilon = 0.00001 with $\sigma = 2$ [as in Dosenbach et al. (2010)]. We use Spider (http://people.kyb.tuebingen.mpg.de/spider/main.html), an object orientated environment for machine learning in Matlab (MATLAB 7.1.0, The Mathworks, Natick, MA), for generating the SVM models.

The brain functional connectivity maturity index (fcMI) for TDC (**Figures 3** and **A1**) was based on the top 300 connections that correlated (and passed FDR correction) in each round of leave-one-out-cross-validation (LOOCV). Calculating fcMI by LOOCV involves removing a single subject as a test sample, then using the remaining data for feature selection and as the training set for the SVR predictor. This procedure is then repeated until each subject is used once as the test case [i.e., univariate feature-filtering, see De Martino et al. (2008); Pereira et al. (2009)].

We also conducted an SVM analysis using features (i.e., connections) that most strongly differentiated ADHD subjects of each of two subtypes [i.e., primary inattentive type (ADHD-I) and combined type (ADHD-C)] from each other and/or control subjects (2-group classification: *t*-test; 3-group classification: ANOVA). For this procedure, we ran the SVM using LOOCV on all subjects conducted on the top 150 features that differentiated each subtype and controls (via the comparisons as noted above). These analyses are presented in **Figure 8** and **Figures A4** and **A5**. The number of features used in this case was reduced because of our reduced sample size for the comparisons.

For the typically developing population, the SVR and univariate age relationships were conducted after 10 potential motion correction procedures (**Table 2**). Procedure 1 used traditional motion correction, which included regression of the BOLD data on 6 rigid body motion parameters (based on the middle reference frame) during preprocessing. Procedure 2 added an additional step that included the six rigid body frame-to-frame parameters, $[\Delta d_{ix}\ \Delta d_{iy}\ \Delta d_{iz}\ \Delta\alpha_i\ \Delta\beta_i\ \Delta\gamma_i]$, as covariates to generate partial correlation matrices (160×160) for each participant. Procedure 3 was similar to Procedure 2, but used FD_i as the covariate. Procedure 4 was the same as Procedure 1, but replaced the traditional preprocessing motion regressors (i.e., Procedure 1) with the frame-to-frame measures (i.e., $[\Delta d_{ix}\ \Delta d_{iy}\ \Delta d_{iz}\ \Delta\alpha_i\ \Delta\beta_i\ \Delta\gamma_i]$). Procedure 5 utilized a group-level correction, similar in nature to what has been proposed by Van Dijk et al. (2012). In this instance, mean FD for each participant is used as a covariate when cross correlating each connection across subject age. Procedure 6 involves a movement "matching" procedure (analogous to matching a parameter between groups) whereby mean FD is used to remove subjects until there is no relationship between mean FD and age (**Figure 5**). Procedure 7 is

Table 2 | Brief description of each motion correction procedure.

Procedure	Method
1	Traditional motion correction: regression of BOLD data with six rigid motion parameter in pre-processing
2	The six rigid body frame to frame parameter were used as a covariates to generate partial correlation matrices
3	Same as Procedure 2 but uses FD as a covariate
4	Same as Procedure 1 but regressed the BOLD data with six rigid body *frame to frame* parameters
5	Group level correction: mean FD is covaried while cross-correlating connections across subject age
6	Movement matching: removed subjects to make sure there was no relation between mean FD and age
7	Frames having FD > 0.2 mm were removed (scrubbing). To account for temporal blurring one frame before and two after were also removed
8	This is combining Procedure 7 and Procedure 6 i.e., frame censoring and movement matching was done
9	Same as Procedure 8. Instead of FD > 0.2, frames were removed based on DVAR > 4, then the movement matching was done with mean DVAR and age
10	Utilizes a polynomial generated for each subject based on information obtained from Procedure 7 to correct for the effect of movement

a method (i.e., "scrubbing" or volume censoring) first proposed by Power et al. (2012a). In this instance, frames are removed based on the magnitude of FD_i. In the original publication frames were removed based on having $FD_i > 0.5$ mm and a DVARS level greater than 5. Here we apply a more stringent criterion based on findings in **Figure 1**, and remove frames prior to creating the 160×160 correlation matrix for each subject based on an $FD_i > 0.2$ mm. To account for potential temporal blurring of the motion-related artifact during the bandpass filter, 2 frames after the censored frame, and one frame before are also removed (Power et al., 2012a). Procedure 8 is an extension of procedure 6 and combines matching with frame removal. Here additional participants are removed to assure there is no relationship of mean FD and age for the *remaining* frames after applying frame removal based on $FD_i > 0.2$. Procedure 9 was identical to Procedure 8, but $DVARS_i$ is used to remove frames instead of FD_i. Here $DVARS_i > 4$ was used to remove frames likely associated with excess movement. In addition, subjects were removed so there was no relationship between mean DVARS and age (see **Figure 5**). Procedure 10 aims to utilize the censoring techniques described above; however, it does so only to generate the effect of movement on the correlation coefficients in each subject. It then utilizes this information to correct the original values. In this sense r-values are corrected via censoring, but without requiring actual data removal for the correlation estimates. For the procedure, connectivity matrices are first generated from the original and censored time series (i.e., Procedure 7) for each subject. Next, within each participant, delta r is calculated for all of the connections (i.e., the difference between the scrubbed and the original connectivity data is calculated for each individual subject). A polynomial curve is fitted to the relationship between delta R and distance between the ROIs (see **Figure A8**). The general form of a polynomial equation is shown below.

$$y = p_1 x + p_2 x^2 + p_3 x^3 + \cdots + p_n x^n + p_{n+1}$$

As shown in the **Figure A8**, for a given subject, the red curve shows the trend of delta r and the blue curve shows the curve fitted to delta r (Note: in the figures we also include LOWESS curves of the data to visualize how the polynomial fits the data). The method is applied on each subject individually such that every subject will have an equation that fits their respective curve. This equation is then used to regress the effect of movement on the original r-values. Censoring the frames as in Procedure 7 after applying this procedure reveals no effect of movement (**Figure A8**).

VISUALIZING REGIONAL BRAIN CONTRIBUTIONS

Regional brain contributions associated with age or ADHD status were not equally distributed throughout the cortex. Some regions had many connections that strongly correlated with age (or differentiated TDC from ADHD subjects), while others had few or none. As such, we represented this phenomenon with a parameter termed "Node Strength" by scaling the diameter of each node by the summed z-scores of each of its connections identified as a connection of interest (COI) in previous analyses. Node strength

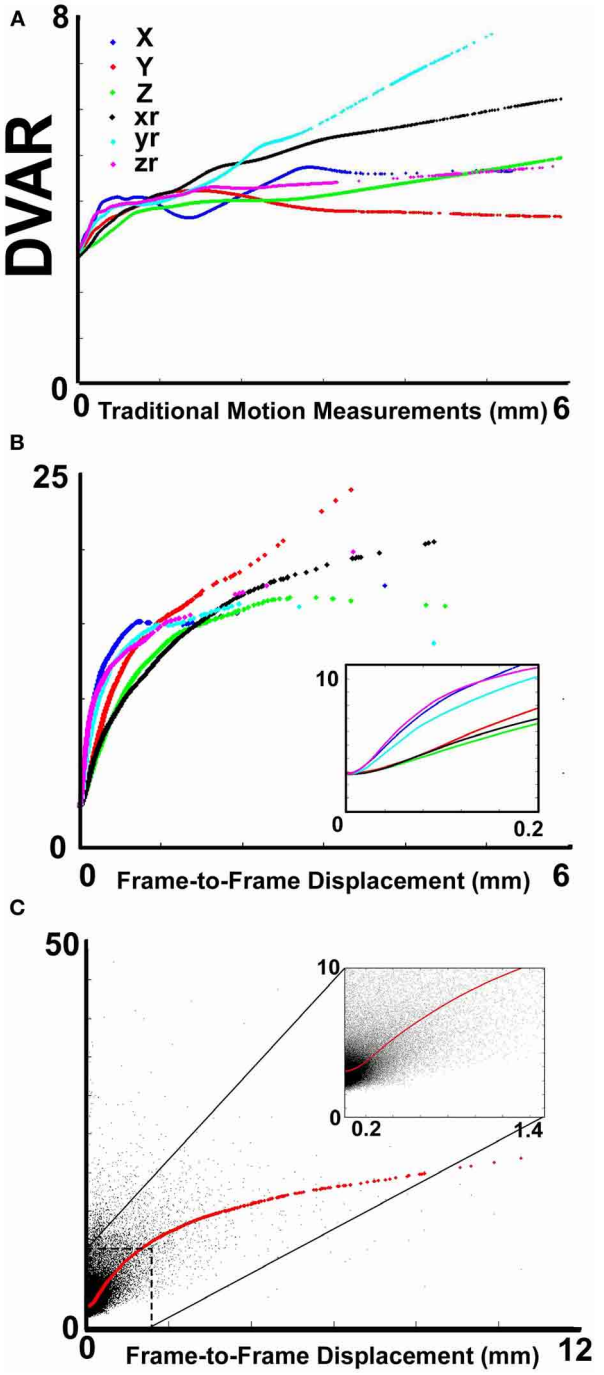

FIGURE 1 | The relationship between movement parameters and variations in echo-planar imaging. (A) LOWESS curves of six traditional movement measurements, which align any given frame within a run to a reference frame within that run, relative to the derivation in BOLD signal (i.e., DVARS) from volume to volume (or frame to frame). These traditional movement measurements show a non-linear relationship with changes in the BOLD signal. **(B)** The same plot as in **(A)**, using 6 frame-to-frame motion measurements as opposed to traditional movement parameters. **(C)** FD, which combines all 6 directions in **(B)**, as a function of DVARS. There is a stronger relationship of DVARS and the FD measurements, compared to traditional motion estimates shown in **(A)**.

is simply the sum of all of the weights (in this case z-scores) attached to a specific node [e.g., see Hagmann et al. (2010)]. Node strength was calculated for the consensus features (those features present in each round of the LOOCV procedures above) that most strongly correlated with age, as well as for the strongest feature set found in the comparisons with regard to the ADHD subtypes. If an ROI did not have any COIs that correlated with age or differentiated groups based on our criterion, it was assigned a Node Strength of 0. Nodes with many and/or strong connections that correlated with age or differentiated groups have a high node strength. Raw values for all Node Strengths are presented in **Table A1**. Surfaced-based mapping of nodes and vectors were conducted using caret software (Van Essen et al., 2001; Van Essen, 2005).

RESULTS

Given prior demonstrations of increased micro-movements in pediatric samples and the potential for motion artifact to contribute to estimates of functional connectivity (Power et al., 2012a; Satterthwaite et al., 2012; Van Dijk et al., 2012), we first provide a comprehensive examination of micro-movements in the ADHD-200 sample, as well as the impact of various correction strategies. Then, we report unique findings from our ADHD-C and ADHD-I samples, after controlling for movement using the most conservative procedure.

SECTION 1: CHARACTERIZING AND ACCOUNTING FOR MICRO-MOVEMENTS IN THE ADHD-200 SAMPLE

We first examined the relationship between traditional movement measures and abrupt changes in the BOLD signal as measured across the whole brain. Traditional movement measurements relate to the six directional adjustments (i.e., rotation and translation) required to align any given frame within a BOLD fMRI run to a reference frame within that run. As noted in methods, this reference frame can be any frame (e.g., in AFNI), but often is the middle frame of a run (e.g., in FSL). Our measurement for changes in the whole brain signal is termed DVARS. **Figure 1A** clearly illustrates that there is a non-linear relationship between all six traditional motion parameters and changes in the BOLD signal. These findings violate the assumptions inherent to several traditional preprocessing movement correction procedures (e.g., linear regression), for both traditional task and connectivity-based fMRI analyses.

Importantly, however, while the relationship of frame-to-frame motion parameters (i.e., FD) and changes in the whole brain signal is positive, there is still a non-linear relationship (**Figure 1B**). These data suggest that simply replacing traditional movement parameters with frame-to-frame motion parameters in traditional preprocessing steps (i.e., linear regression) will assist, but likely not fully account for potential movement artifact. This presumption was confirmed in the following analyses (also see **Figure A1**).

Traditional motion correction in a typically developing sample (Procedure 1)

We begin this section by first identifying the functional connections that most strongly associated with age in TDC, when only traditional motion correction procedures were used. Again, for each subject, we estimated the Pearson *r* between the resting BOLD time-series computed for 160 *a priori* defined brain ROIs (Dosenbach et al., 2010). After Fisher r-to-z transformation, this step created 455 square *r*-matrices (160 × 160)—containing 12,720 pair-wise connections for each participant. Two analyses were applied to examine age relationships. The first was a basic cross-correlation to examine how each pair-wise connection relates with age. The second was the use of SVR-based MVPA (Norman et al., 2006; Rizk-Jackson et al., 2011) to verify the ability of rs-fcMRI to predict brain maturity for TDC. Chronological age served as our training measure, while predicted age was our measure of functional maturity. We calculated the fcMI by setting the model fit (Von Bertalanffy's growth curve model $[a\bullet(1 - e^{-bx})]$) of predicted age for the oldest subjects equal to one. Predictions for each TDC subject were based on the top 300 connections in each round of LOOCV.

Consistent with prior reports (Fair et al., 2007a, 2009; Kelly et al., 2009; Supekar et al., 2009; Dosenbach et al., 2010), we found that connections that became stronger with increasing age, or "grew-up," tended to link distant regions (i.e., long-range connections). Connections that became weaker with increasing age, or "grew-down," were primarily found between proximal regions (i.e., short-range connections) (see **Figures 2** and **4**; $p < 0.0001$). We also saw robust age-predictions for the TDC group. Von Bertalanffy's growth curve model ($[a\bullet(1 - e^{-bx})]$) provided a significant fit for predicting brain maturity (**Figure 3A**) ($R^2 = 0.36$). These results served as our baseline.

Using frame-to-frame measures as covariates when correlating time series (Procedures 2–4)

We next examined the same age relationships using the additional (or alternative) procedures focused on micro-movement correction. Here, in addition to performing the traditional correction for motion (Procedure 1), we computed partial correlation estimates between the BOLD time series of each of the 160 ROIs using either (1) Frame-to-frame displacement (FD_i) as a covariate, or (2) the frame-to-frame 6 motion parameters as unique covariates (i.e., $[\Delta d_{ix}\ \Delta d_{iy}\ \Delta d_{iz}\ \Delta\alpha_i\ \Delta\beta_i\ \Delta\gamma_i]$). This step created 455 square *partial* correlation matrices for each method. Again, after Fisher r-to-z transformation, each pair-wise connection was cross-correlated with age. As can be observed in **Figures A1** and **4**, the distinction between short- and long-range connections in terms of their relationships with age was only modestly altered by these additional correction procedures, while the ability to make age predictions was maintained. Similar findings were observed when we replaced the traditional movement regressors in the preprocessing with regressors based on FD_i (**Figures A1** and **4**). The limited effect of these procedures on the short-long range distinctions is likely secondary to the non-linear relationships between changes in BOLD and movement (**Figure 1**).

Using mean FD as a covariate with age (Procedure 5)

Similar to prior work by Van Dijk et al. (2010), we next examined the effect of including mean FD_i across all frames for each subject as a group-level covariate in the calculation of the correlation

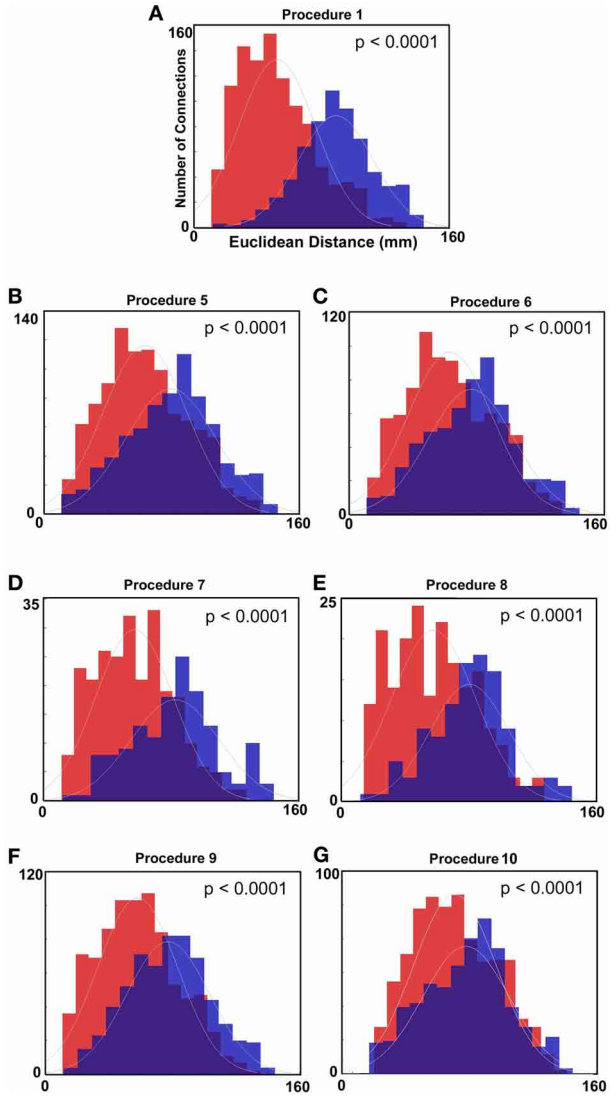

FIGURE 2 | Histograms of Euclidean distance for functional connections that get stronger with age and those that get weaker with age (FDR corrected). Red colors denote the distribution with regard to distance of those connections that get weaker with age. Blue colors represent the distribution for those connections that get stronger with age. **(A)** Distance measurements after traditional motion correction (Procedure 1), **(B)** Procedure 5, **(C)** Procedure 6, **(D)** Procedure 7, **(E)** Procedure 8, **(F)** Procedure 9, and **(G)** Procedure 10. **Table 2** provides a short summary of each of these procedures. Each motion correction procedure reduced the mean difference in distance between those connections that get stronger relative to those that get weaker with age. Also see **Figure A1** for procedures 2–4.

between functional connectivity and age. As previously shown, this method substantially attenuates the distinction between long- and short-range connections in terms of their relationships with age (Van Dijk et al., 2012) (see **Figures 2** and **4**). The difference in the mean distance of short-range connections using this method relative to those obtained using Procedure 1 (i.e., traditional correction) was highly significant ($p < 0.0001$). The difference in mean distance of long-range connections using this method relative to Procedure 1 was also highly significant ($p < 0.0001$) (**Figure 4**). Interestingly, while the short-long range effect was reduced, the ability to make age predictions in the data rose slightly from the original estimate (i.e. Procedure 1) with the movement correction procedures ($R^2 = 0.39$).

Subject "matching" based on mean FD$_i$ (Procedure 6)

We next employed a method that excluded subjects on the basis of their mean FD until there was no relationship between the measure and age ($p = 0.63$) as shown in **Figure 5C**. Because there appears to be no strongly discernible relationships between the effect of movement on the BOLD signal and age (i.e., the effect of FD on DVARS is similar for all ages—**Figure 5**), we expected this analysis to yield similar results as using the measure as a covariate with regard to our outcome measure (albeit weaker overall significance levels due to the decrease in sample size). For this analysis we removed participants ($N = 76$) until there was no relationship between mean FD$_i$ and age (see **Figure 5C**). We then re-ran the age cross-correlation. This analysis also attenuated the distinction between short- and long-range connections in terms of their relationships with age (**Figures 2** and **4**). The difference in mean distance of short-range connections using this method relative to no "micro-movement" motion correction (i.e., Procedure 1) was also highly significant ($p < 0.0001$). The difference in mean distance of long-range connections using this method relative to no "micro-movement" motion correction was significant, as well ($p < 0.0001$) (**Figure 4**). The ability to make age predictions increased ($R^2 = 0.42$) (**Figure 3**).

Volume censoring (Procedure 7)

Another method recently employed to correct for motion simply censors or removes volumes or frames corresponding to excessive movement (Power et al., 2012a). We used FD$_i$ > 0.2 mm displacement as our threshold for removal of frames. The basis of this threshold was rooted in **Figure 1**, which highlights a knee in the LOWESS curve at, or just below, ~0.2 mm FD. Updated recommendations by (Power et al., 2012b) independently established the same threshold. While age-relationships were not completely removed, we once again see a reduction in the distinction noted between short- and long-range connectivity (**Figures 2** and **4**). The ability to make age predictions based on SVR-based MVPA was similar to Procedures 5 and 6 ($R^2 = 0.42$).

One concern of this approach is the consequence of missing, or removing, values in time-series analysis. We attempted to examine this influence with a simulation. The simulation began by selecting 3 participants with little to no movement based on the 0.2 mm threshold (i.e., 0–1% of frames above the threshold) from each dataset (i.e., each institution). For each of these subjects we than calculated their correlation matrix across the 160 ROIs, which served as our baseline (i.e., the true correlation structure). We then iteratively removed frames from the time courses of these subjects based on the frames that did not pass the 0.2 mm displacement criteria for each of the other subjects in their cohort (based on FD$_i$). In other words if *subject 1* had 30 frames that did not pass criteria, we would remove those same 30 frames from the 3 participants with little to no movement and then recalculate

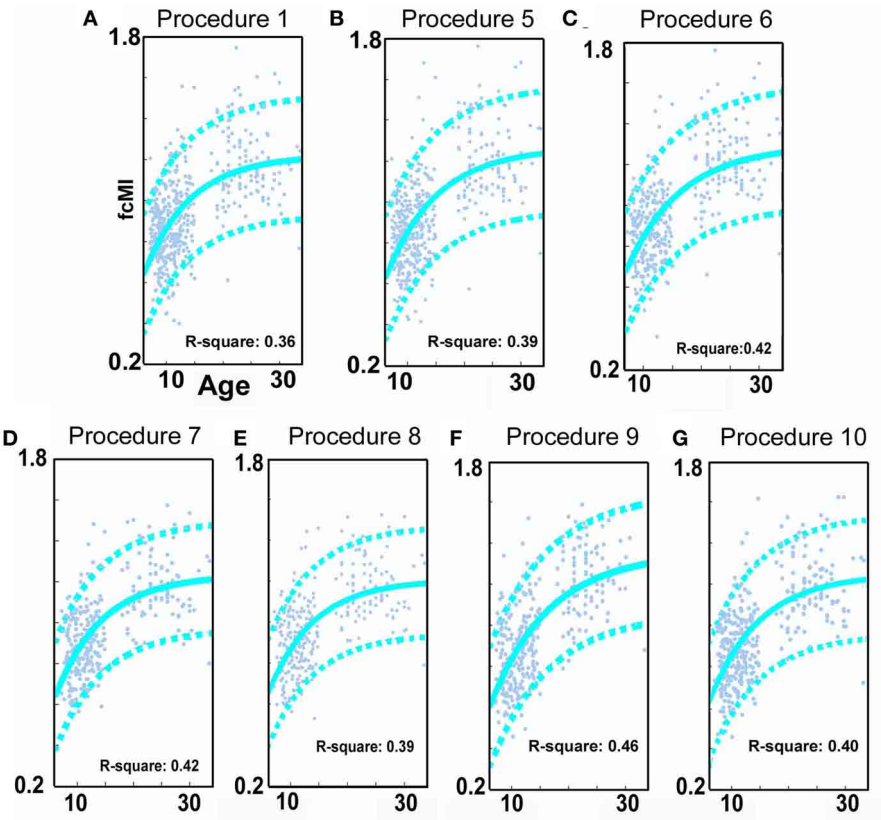

FIGURE 3 | SVR-based MVPA brain maturation curves. SVR-based MVPA brain maturation curves showing individual age predictions for TDC (blue) after seven motion correction procedures. Chronological age is presented on the x-axis and fcMI on the y-axis. The Von Bertalanffy's growth equation shows a reasonable fit for these data (with 90% prediction limits shown with dashed lines). **(A)** Procedure 1 (traditional motion correction), **(B)** Procedure 5, **(C)** Procedure 6, **(D)** Procedure 7, **(E)** Procedure 8, **(F)** Procedure 9, and **(G)** Procedure 10. Also, see Methods and **Table 2** for short description of each procedure.

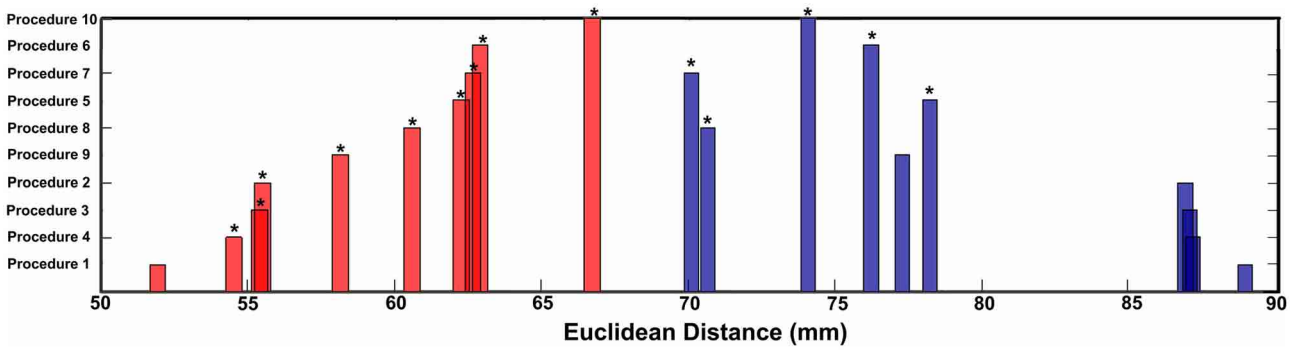

FIGURE 4 | Mean Euclidean distance for short-range and long-range connection changes over age for each procedure. Asterisks represent significant differences in distance relative to Procedure 1 (traditional motion correction).

the correlation matrices for those 3 participants. The procedure was next repeated for subject 2, and so on. In each instance, we were able to compare the newly generated matrices to the baseline matrices of the 3 subjects that had no frames removed. Thus, the simulation gave us a pure reference by which we could systematically quantify the effect of removing a given percentage of frames on the overall correlation matrix's structure. Importantly, the number, percentage, and frequency of frames removed for each iteration conformed directly to real data. **Figure 6** shows the distance plots (i.e., 1–r) of the original, baseline, matrices relative to the simulated matrices as a function of the percentage of frames removed. As can be seen, the distance is low when few frames are

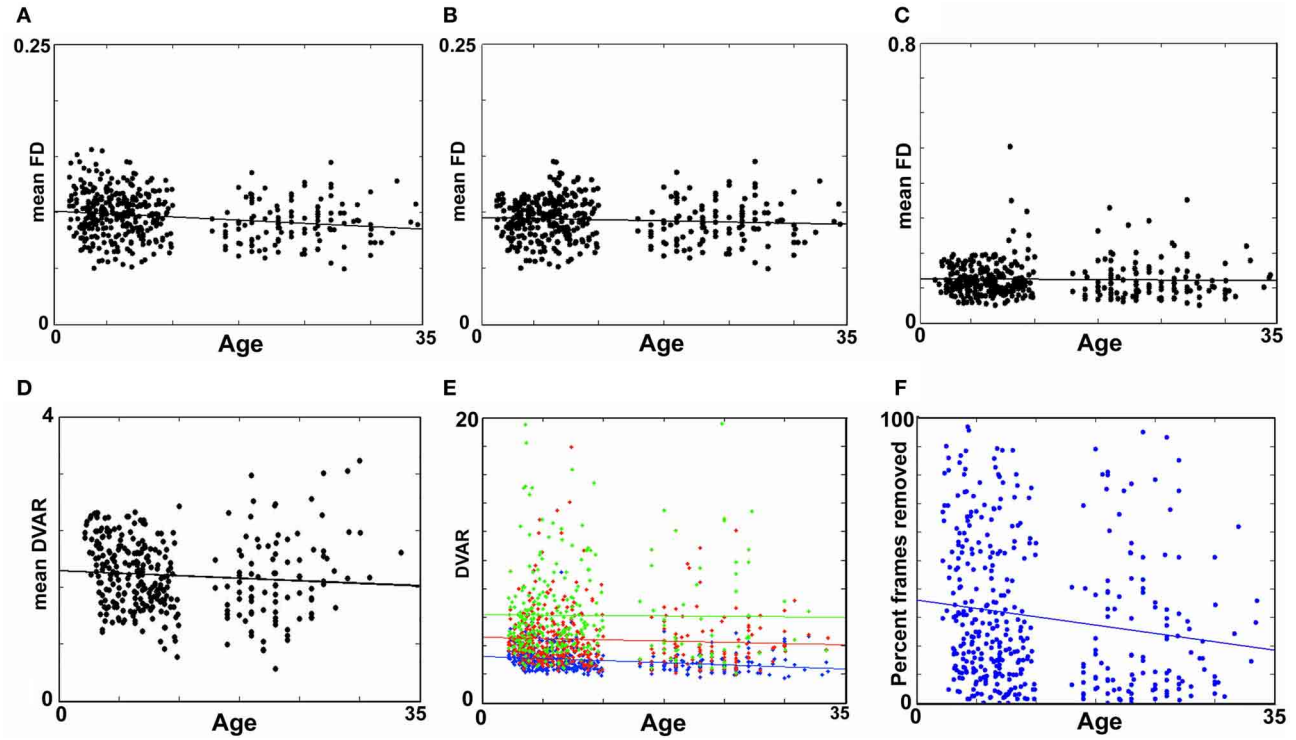

FIGURE 5 | Age relationships of movement-related measurements. **(A)** Mean FD as a function of age for Procedure 7 (i.e., mean FD for the remaining frames after removal). There is a small, but significant relationship with mean FD and age ($p < 0.01$). **(B)** Procedure 8 after matching shows that the relationship found in **(A)** is no longer significant ($p = 0.13$). **(C)** Mean FD as a function of age for Procedure 6 (matching only without frame removal). Again, here there is no relationship with mean FD and age ($p = 0.62$). **(D)** A similar figure as in **(B)**, for the DVAR based frame removal (i.e., Procedure 9). As with **(B)** and **(C)**, there is no age relationship with the DVAR measurement ($p = 0.13$). **(E)** Using mean FD as a covariate (i.e., Procedure 5) with age assumes that movement affects the BOLD signal similarly for any age group. Thus, for all subjects, data frames were split into low (FD = 0–0.2), medium (FD = 0.2–0.4), and high (FD = 0.4–0.6) movement for each participant. The age relationship with our whole brain BOLD measurement, DVARS, was then examined for each bin. While there is a limited age relationship in the lowest movements overall the data suggest movement affects the BOLD signal similarly across the age range studied here (blue $p < 0.01$, red $p = 0.25$, green $p = 0.88$). **(F)** The percentage of frames removed across age for Procedure 8 (i.e., FD frame censoring after matching), as expected shows a slightly higher rate of frame removal for the younger participants.

removed; however, the tight relationship degrades as the fraction of frames removed approaches 1. We therefore re-ran our analysis for Procedure 7 while including only those subjects with <60% of their frames removed. Results were largely convergent (see **Figure A6**).

We also attempted to examine the possibility of replacing missing frames via a common spline interpolation in our simulation [similar in concept to Carp (2011)]. As can be seen in **Figure 6**, imputing missing frames using cubic spline interpolation added significantly more noise than simply removing the volumes of high movement. As such, we did not attempt to interpolate frames for any of the following analyses.

Volume censoring with subject "matching" using mean FD_i (Procedure 8)

Importantly, despite the removal of motion related frames exceeding a threshold of 0.2 mm displacement, age, and mean FD continued to be associated, $p = 0.0001$ (**Figure 5A**). This observation is likely due to a small increase in movement in younger subjects for the remaining frames and what can be observed as a small relationship of frame-to-frame changes in BOLD (i.e., DVARS) and FD even under this 0.2 mm cutoff (**Figure 1**). As such, we repeated the age correlation after combining this censoring approach (Procedure 7) with mean FD matching (Procedure 6—as in **Figure 5B**). Thus, we repeated the analysis after removing participants ($N = 142$) until the there was no relationship of mean FD_i (for the remaining frames) and age ($p = 0.13$). This combined method also reduces the short- and long-range relationships (**Figures 2** and **4**), but shows a slight decline in the ability to make age classifications in individuals ($R^2 = 0.34$). [Note: findings were consistent across individual sites (**Figure A2**)].

Volume censoring with subject matching using mean DVARS (Procedure 9)

Here we examine the utility of using the DVARS measure itself to remove frames. In this instance, instead of removing frames based on FD_i, we removed frames based on changes in $DVARS_i$ ($DVARS_i > 4$). We also removed participants ($N = 130$) to assure no relationship with DVARS and age for the remaining frames. This analysis is shown in **Figures 2**, **4**, and **5D**. Interestingly, while the short- and long-range distinctions in terms of their relationship with age were also significantly reduced

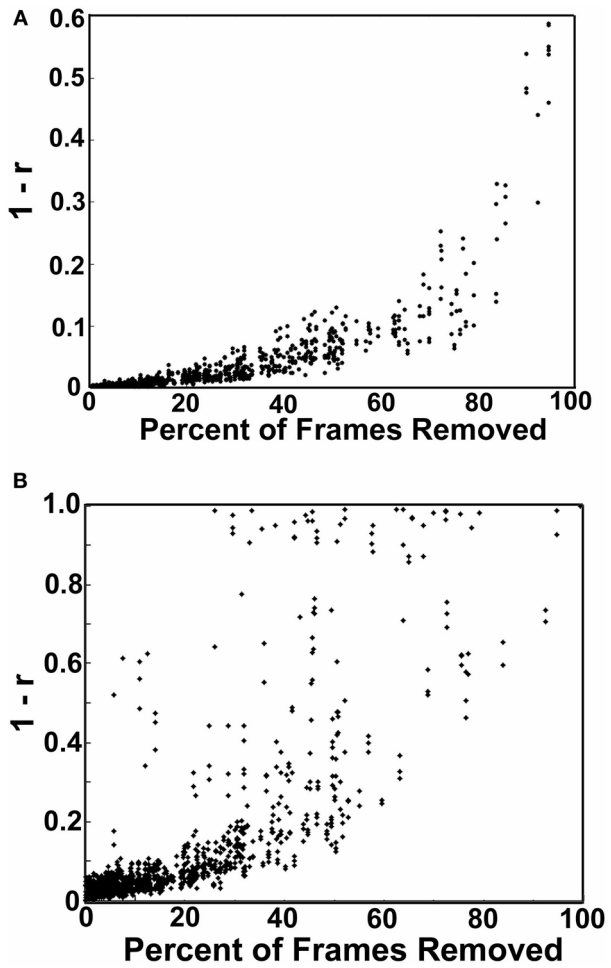

FIGURE 6 | Simulations showing the effect of frame removal on the correlation structure of a subject matrix. The simulation began by selected 3 participants with little to no movement based on the 0.2 mm FD threshold (i.e., 0–1% of frames) from each datasets used here (i.e., each institution). For each of these subjects we then calculated their correlation matrix across the 160 ROIs, which served as our baseline. We then iteratively removed frames from the time courses of these subjects based on the frames that did not pass the 0.2 mm displacement criteria for each of the other subjects in their cohort (based on FD_i). In each instance, we compared the newly generated matrices to the baseline matrices of the three subjects that had no frames removed. In **(A)** we show the distance plots (i.e., 1–r) of the original, baseline matrices relative to the simulated matrices as a function of the percentage of frames removed. The distance is low when few frames are removed; however, the tight relationship degrades as the fraction of frames removed approaches 1. In **(B)** we show the same procedure except in this instance missing frame are imputed via a cubic spline interpolation. Substantially more noise is added by interpolation.

as with other methods, the ability to make age predictions was the highest of all of the methods thus far ($R^2 = 0.46$) (**Figure 3**).

R-value correction using polynomial regression based on volume censoring (Procedure 10)

Finally, we examine a method that applies an r-value calibration by utilizing a polynomial regression based on the correction induced by data scrubbing, as outlined in the methods. Importantly, this particular method is not completely devoid of the issues noted above for data censoring. To obtain a valid estimate of the true polynomial, the method still requires a sensible number of remaining frames. As such, we ran the analysis on subjects with at least 40% of their frames remaining after censoring. The analysis is shown in **Figures 2**, **4**, and **5**. As with the other procedures, the short- and long-range distinctions in terms of age were significantly reduced, yet the ability to make age predictions remained ($R^2 = 0.40$; **Figure 3**).

In **Figure 7** (also see **Figure A3**), for this procedure along with procedures 5 and 8, we use node strength, a commonly employed graph-theory metric (see Methods), to reveal the nodes that are most strongly predicative of age. The figure shows a distributed pattern of connectivity with several subcortical structures being highly predictive. Connections that tended to get stronger with age appeared to be within networks, and connections that got weaker with age were between networks. The findings were qualitatively similar across the motion correction methods.

SECTION 2: PREDICTION OF DIAGNOSTIC STATUS

Consistent with prior work (Dosenbach et al., 2010), we followed our movement analyses with an effort to determine if rs-fcMRI in combination with SVM-based MVPA is capable of predicting ADHD status in children with ADHD-C or ADHD-I. Predictions for each subject were based on the top 150 connections (features) that differentiated each subtype from controls in each round of LOOCV. To avoid biasing our classifier group, subsample size was matched (see **Table 3**). Consensus features, or those connections present across all of the LOOCV samples are provided in **Figure A5**. In addition, node strength was used to further examine the functional neurobiology that differentiated the two ADHD subtypes from controls. Raw absolute values of node strength are provided in **Table A1**.

For prediction analyses and subtype comparisons we used what we saw as the three most conservative movement correction approaches, Procedures 5, 8, and 10. These are not necessarily "the best" procedures, as performance varied based on whether we examined short- or long-range connections. Importantly, these analyses were careful to consider potentially confounding variables (site, gender, and IQ) by controlling for them using multiple regression analysis. That is, these variable connections were regressed from functional connectivity values across ROIs connections of all subjects. When adjusting for the confounding variables, our regression equation is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon,$$

where y represents the vector of connectivity values, β_0 is the intercept, ε is the vector of residuals, β_1 to β_4 represent regression coefficients corresponding to independent or explanatory variables: x_1—ADHD Diagnosis, x_2—Site, x_3—Gender, and x_4—IQ. The unknown β_i's that measure the sample relationship between y and x_i after all other variables have been partialled out, are obtained using ordinary least

squares (OLS) estimation. Using the OLS estimates of β_i, functional connectivity values adjusted for confounding variables are given by:

$$y = \beta_0 + \beta_1 x_1 + \varepsilon.$$

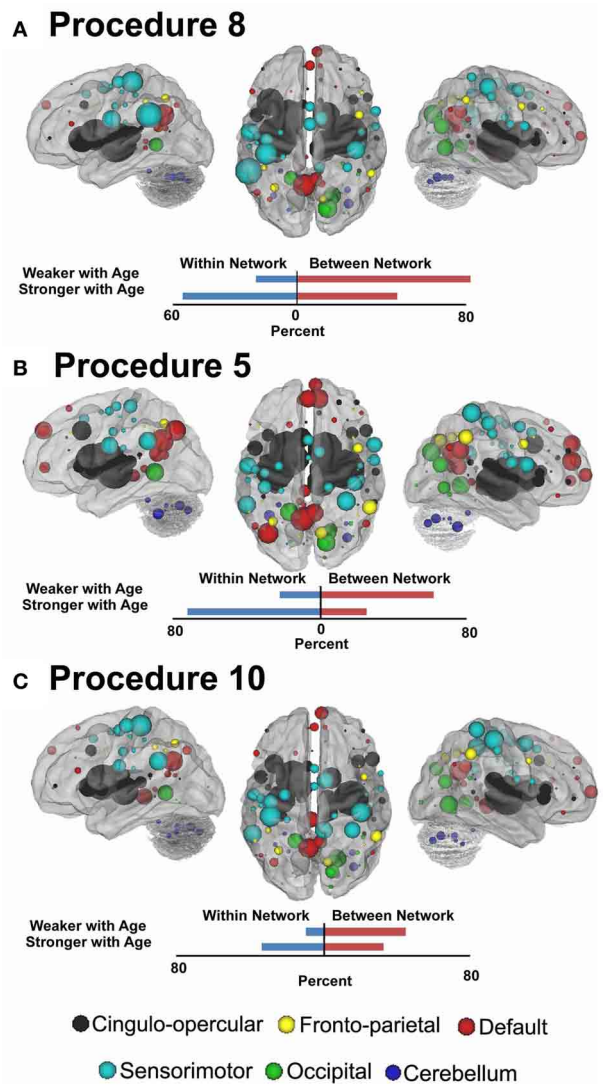

FIGURE 7 | Node strength for the consensus features that best predict age for Procedure 8 in (A), Procedure 5 in (B), and Procedure 10 in (C). Within and between network comparisons are also presented for those connections that get stronger with age vs. those that get weaker with age.

Using procedure 5, our predictions for ADHD-C relative to controls showed 77.0% accuracy (75% sensitivity and 76.9% specificity), while our predictions for ADHD-I relative to controls showed 80.8% accuracy (78.9% sensitivity and 82.7% specificity). Using procedure 8, our predictions for ADHD-C relative to controls showed 63.4% accuracy (61.5% sensitivity, 65.4% specificity), while our predictions for ADHD-I relative to controls showed 78.8 % accuracy (75.0% sensitivity, 82.7% specificity). Last, using procedure 10, our predictions for ADHD-C relative to controls showed 71.2% accuracy (73.1% sensitivity, 69.2% specificity), while our predictions for ADHD-I relative to controls showed 82.7% accuracy (78.9% sensitivity, 86.5% specificity).

We also attempted a 3-group classification for which chance is 33%. For procedure 5, the 3-group classifier revealed results that were highly significant (overall 69.2% accuracy; TDC = 67.3%, ADHD-C = 71.1%, ADHD-I = 69.2%). For procedure 8, the results were also promising albeit reduced (overall 56.4%; TDC = 59.6%; ADHD-C = 44.2%; ADHD-I = 65.4%). For procedure 10 the findings again revealed largely consistent strong findings (overall 68.6%; TDC = 71.2%, ADHD-C = 63.5%, ADHD-I = 71.2%).

Characterizing the neural correlates of ADHD subtypes

To examine the potential neurobiological differences in subjects with inattentive and hyperactive symptoms (i.e., combined type), and those with inattentive symptoms only (i.e., predominantly inattentive type), we examined the features that most strongly differentiated the groups for each of the movement correction procedures. That is, we calculated *node strength* for each group comparison as described in methods.

For procedure 10, visualization of those nodes whose strength most strongly separated each of the two ADHD subtypes considered from TDC (i.e., Node Strength), revealed stark differences between the two (**Figure 8**; **Figures A4**, **A5**; **Table A1**). Nodes most strongly differentiating ADHD-C and TDC were distributed across the cortex (**Figure 8**), but were most prominent in the medial prefrontal and posterior parietal nodes of the default network. Other atypical regions included nodes of the sensorimotor, visual, and cingulo-opercular systems. Findings with regard to consensus features and ADHD-I were similar to those of ADHD-C in that they were largely distributed across systems including prominent features in the sensorimotor systems; however, the patterns between the subtypes were also quite distinct. In the case of ADHD-I, atypical nodes in left and right dorsolateral prefrontal regions along with the cerebellum were more strongly predictive. These particular nodes (see **Figure A6**) have been empirically linked with community detection methods to fronto-parietal systems (black/yellow) (Dosenbach et al., 2010).

Table 3 | Subject characteristics for ADHD and control comparisons.

		N	Age (mean)	% Female	IQ	FD mean (pre-censor)	FD mean (post-censor)
Total	TDC	52	11.12	0.48	115.83	0.12	0.10
	ADHD-I	52	11.63	0.23	108.27	0.14	0.10
	ADHD-C	52	10.98	0.17	109.04	0.13	0.10

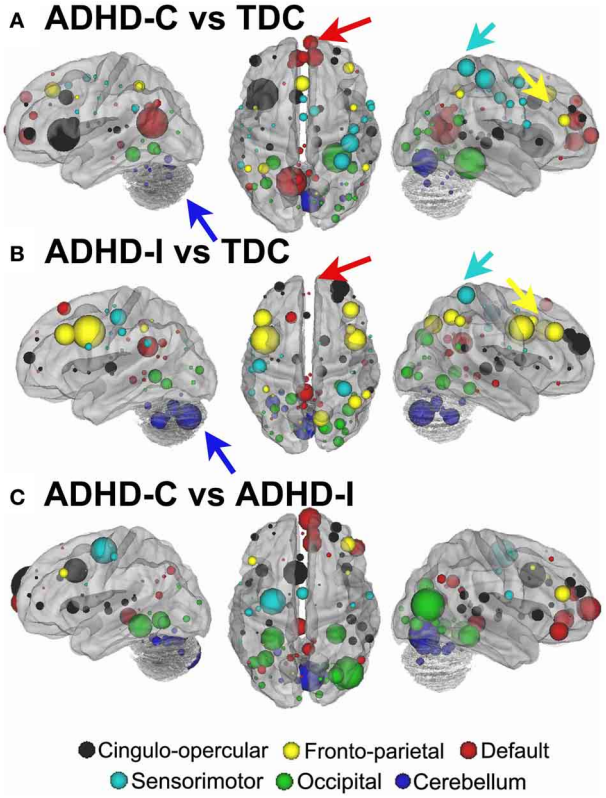

FIGURE 8 | *Post-hoc* comparisons on the consensus features from the 2-group classification for Procedure 10 (see SI Figure 4 for Procedures 5 and 8). The two most common subtypes of ADHD (ADHD-I and ADHD-C) show distributed patterns of atypical connectivity relative to TDC, as measured with node strength. **(A)** Node strength for TDC vs. ADHD-C shows strong differentiation in regions of the medial prefrontal cortex among other distributed systems. **(B)** Node strength for TDC vs. ADHD-I shows differentiation distributed throughout the cortex as well, with prominent nodes including bilateral dorsolateral prefrontal, and cerebellar regions among others. **(C)** Comparisons between the subtypes show similar trends. [Node colors represent network categorization from a community detection procedure performed for a previous report (Dosenbach et al., 2010). Red—default; blue—cerebellum; yellow—fronto-parietal; black—cingulo-opercular; green—occipital; cyan—sensorimotor].

In addition, similar to ADHD-C, sensorimotor regions (light blue) (**Figure 8**) were also atypical in ADHD-I. The direct comparison between ADHD-I and ADHD-C were largely consistent with the distinctions noted above (**Figure 8**). As can be seen in **Figure A4**, procedures 8 and 10 revealed largely consistent findings with regard to overall patterns.

DISCUSSION

Consistent with recent work demonstrating the utility of data aggregation (Biswal et al., 2010), and in the spirit of the ADHD-200 Global Competition (ADHD-200-Consortium, 2012), we used the large-scale aggregate ADHD-200 dataset to highlight the need for stringent micro-movement correction. The work also provides the first insights into common and distinct neural substrates of the two most common ADHD subtypes. Application of graph-theoretic metrics implicated atypical connectivity across multiple systems in the combined subtype (ADHD-C), particularly in the default network. Although similar observations were present in the predominantly inattentive subtype (ADHD-I) functional connectivity differences were more evident in bilateral dorsolateral prefrontal regions, and in the cerebellum, rather than in the default mode network. Classification analyses aimed at predicting individual diagnostic status based on patterns of functional connectivity further highlight differential neural substrates for the two subtypes, and serve to provide a vision for future clinical applications of neuroimaging.

MOTION CORRECTION IMPROVES ADHD CHARACTERIZATION

There were several notable findings regarding motion correction. For example, we showed that the relationship between movement measurements based on traditional reference frame calculations can, in some instances, be misleading. In particular, the relationship between traditional movement measurements and the BOLD signal is markedly non-linear. As such, traditional functional connectivity and fMRI movement correction that utilize linear regression are likely to be limited with regard to micro-movement correction. Importantly, however, while frame-to-frame measurements (i.e., FD), appear to be superior with relation to the assumption of linearity, they too share a non-linear relationship with the BOLD signal. Thus, simple linear regression of these measurements from the BOLD signal also provides only limited improvement with regard to motion artifacts (**Figure 4**). With that said, we found several methods that appeared relatively robust at controlling for motion artifact in the BOLD signal. All of these additional approaches reduced the distinction between short- and long-range connections in terms of their relationships with age.

Age-related growth curves were improved with several of the motion correction procedures, with the strongest age-related effects occurring when using DVARS and matching as the surrogate measure for frame removal (**Figure 3**). Nevertheless, there are some concerns with approaches that require frame removal. In particular, the approaches reduce sample size due to "matching" (i.e., subject exclusion) or, effectively removing subjects due to the high percentage of frames removed. Thus, statistical power is reduced. The large sample size here afforded us the ability to examine these conservative approaches, but other samples may not support the same opportunity.

Another concern relates to the potentially undesirable effects of frame removal on time-series analyses. Our simulations suggest that the correlation structure can be altered with frame removal; however, the concern appears to be largest only after significant portions of the frames are removed. One potential way to overcome this limitation for the frame removal procedures would be to interpolate missing values. While in theory we agree with this suggestion, our simulations suggest that the introduction of interpolation as a means of minimizing data-loss can be counter-productive, unless more elaborate approaches are developed and validated beyond the common methods attempted here [e.g., Carp (2011)]. Another alternative to avoid frame removal is provided with Procedure 10. This procedure utilized censoring

techniques to generate a polynomial capable of correcting the original *r*-values, and was quite productive at motion correction. The results were promising, but the method is not able to completely resolve the issues noted by frame removal. The method still relies on a valid estimate of the correlation structure after removing frames, and thus still requires participant removal when a high degree of motion is present.

The use of mean movement parameter (i.e., mean FD) as a covariate in the group analyses also appeared to be a productive method to reduce motion-related artifact in the BOLD signal. This procedure was similar to that originally proposed by Van Dijk et al. (2012). The results using this procedure also show a significant reduction in short- and long-range indices, albeit less than frame censoring methods (see **Figures 2** and **4**). Importantly, the method does not suffer from the inherent reduction of power or generalizability from decreasing sample size.

Nonetheless, the three techniques reveal similar, albeit not identical findings. It should be emphasized, however, that while these approaches appear to perform respectably for reducing artifact induced by changes in head position, work in this area is currently ongoing and even more robust procedures are likely to emerge.

Similar to a recent report (Satterthwaite et al., 2012), our findings show that many of the developmental principles identified previously are still present even after stringent motion correction, while others are not. For example, age-related predictions are still possible and were even improved with motion correction procedures (Dosenbach et al., 2010). However, while the changes in short- and long-range connectivity appear to remain, this phenomena is significantly reduced after motion correction procedures (and becomes only weakly significant for some procedures)—highlighting that the distance between those connections that get weaker with age and those that get stronger overlap more than was previously appreciated. Network configuration, while not identical, appears to be similar to that of adults, even in our younger ages, as shown in **Figure A7** and in Power et al. (2012a). Thus, while significant changes do occur throughout development, the general network structure appears to be intact at early ages. It is possible that similar findings will unfold for the aging literature as well (Meunier et al., 2009; Van Dijk et al., 2012). Most importantly, however, these data in aggregate highlight the need to take careful consideration of movement related artifact in all connectivity studies, particularly during classification of clinical samples as noted below.

THE FUNCTIONAL NEUROBIOLOGY OF ADHD-C AND ADHD-I IS DISTINCT

Over the course of the last decade, the validity of DSM-IV subtype classifications has been repeatedly challenged. Combining statistical analysis with a variety of neuropsychological, behavioral and clinical observation approaches, some have questioned the validity of distinguishing between combined and inattentive subtypes, while others have called for considerations of further subgroupings [for a review see Willcutt et al. (2012)]. Until now, neuroimaging studies have contributed little to this debate—largely for lack of sufficient data. Using rs-fcMRI, we found that while commonalities exist in the neural substrates of the two subtypes, distinctions are also present—suggesting that ADHD-I may capture not simply a "less severe" form of ADHD-C, but a neurobiologically distinct syndrome. This reinforces the conclusion from Willcutt et al. (2012) that additional neuroimaging data on the question of ADHD subtypes is sorely needed. It will remain to be seen whether longitudinal brain imaging data can help resolve the problem of subtype instability (Lahey and Willcutt, 2010) in the disorder.

Consistent with recent work emphasizing the value of network-based approaches for understanding psychiatric illnesses such as ADHD (Castellanos et al., 2008; Fair et al., 2010; Costa Dias et al., 2012; Mills et al., 2012), observed differences among the subtypes can be understood in terms of their network distributions. Findings for both subtypes were relatively distributed, with each affecting only a subset of regions in any one network. However, most notable in the ADHD-I group was the prominence of differences in specific regions of the fronto-parietal and cerebellar systems—specifically dorsolateral prefrontal regions. The fronto-parietal network is believed to be an important system for task level control (Dosenbach et al., 2008). Given that ADHD-I subtype symptoms are predominantly those of inattention, this finding is largely consistent with the idea that this subtype is characterized mainly by problems in task level control systems. While both subtypes revealed largely distributed atypical connectivity, findings for the ADHD-C were more prominent in the default network. These findings for ADHD-C are consistent with a growing body of studies implicating the default network in ADHD (Castellanos et al., 2008; Fassbender et al., 2009; Fair et al., 2010), and is intriguing given a growing number of studies highlighting the motivational as well as affective aspects of the disorder (Sagvolden et al., 2005; Musser et al., 2011) [for a reviews see Nigg and Casey (2005); Castellanos et al. (2006)]. In particular, impairments in incentive salience, motivation, and reward processing are increasingly appreciated in ADHD. While speculative, it should be noted that key components of the functional neuroanatomy of these processes are based within the default network (e.g., ventral striatum, ventromedial cortex). Recent theories suggest the default network is associated with remembering the past, as well as planning for, and anticipating future events (Buckner and Carroll, 2007; Buckner et al., 2008). As stated by Buckner and colleagues, the default network may support a "set of processes by which mental simulations are used adaptively to imagine events beyond those that emerge from the immediate environment." Analogous to this notion, the nature of "motivational"-based theories is that children with ADHD are unable to correctly explore, anticipate, and "value" outcomes between present action and future rewards. If the default network is indeed important for using past experiences to explore and anticipate future events, one might anticipate that aberrations in this network and its links to other systems would map onto the ADHD phenotype and, in particular, the impulsive and hyperactive behavior of the ADHD-C subtype, as found here. Studies of gray matter thickness (Shaw et al., 2007) and connectivity measurements (Shannon et al., 2011) have shown that many of these same regions correspond to various forms of impulsivity measures, supporting this claim.

Along with the other systems noted above, ADHD-associated differences in connectivity of primary sensorimotor systems was observed across both subtypes. Children with ADHD often demonstrate difficulties with motor control, paralleling difficulties with higher-order executive control (Pennington and Ozonoff, 1996; Mostofsky et al., 2003). A consistent characteristic of children with ADHD is that they fail to meet age-norms on timed repetitive and sequential movements and manifest a greater amount of motor overflow than age-matched controls (Denckla and Rudel, 1978; Mostofsky et al., 2003). Furthermore, in parallel with numerous recent studies showing ADHD-associated increases in intrasubject response time variability (Castellanos et al., 2005; Vaurio et al., 2009; Epstein et al., 2011), children with ADHD show increased variability during performance of simple motor skills and during motor adaptation (Izawa et al., 2012). Both overflow and variability in motor execution have been found to be related to ADHD-associated impairments at the behavioral level (Mostofsky et al., 2003). Converging evidence from behavioral, imaging, and electrophysiologic studies suggest that ADHD is associated with dysfunction across multiple parallel frontal-subcortical circuits and while it may not be the case that ADHD is a primary motor disorder, it does seems likely that core ADHD impairments are reflected in motor function in practically measurable and biologically meaningful ways (Mostofsky and Simmonds, 2008). To that end, recent studies of motor cortex physiology using Transcranial Magnetic Stimulation (TMS) reveal that subjects with ADHD show reduced Short Interval Cortical Inhibition (SICI) and that lower SICI is robustly correlated with parent ratings of more severe hyperactive/impulsive and inattentive behavior (Gilbert et al., 2006). The findings from the current study further suggest that patterns of sensorimotor connectivity may prove effective in identifying ADHD, particularly those with the combined subtype.

In summary, our findings demonstrate the potential utility of functional MRI approaches in characterizing clinical heterogeneity in ADHD, which should be similarly useful with other psychiatric disorders. While the present work focused on testing the existing categorization among individuals with ADHD, the previous work highlighted above and the distributed nature of our current findings is likely suggestive of significantly more heterogeneity in this population. Future work is likely to benefit from stronger, perhaps more homogenous subgroupings via the application of data-driven methods that identify clinical subtypes based upon variation in the patterns of neuropsychological measures (Fair et al., 2012) or brain connectivity itself. Additionally, given clinical observations that children with ADHD-C over the course of development later present as ADHD-I based on currend diagnostic measures, future work may benefit from comparison of "converters" with those who were diagnosed with ADHD-I throughout their childhood. Such endeavors will undoubtedly require the acquisition of large and tightly coordinated data acquisition across centers.

TOWARD IMAGING-BASED CLASSIFICATION OF DIAGNOSTIC STATUS

Beyond a simple characterization of the neural signatures of ADHD subtypes, the findings presented here highlight the translational potential of rs-fcMRI for developmental neuropsychiatric disorders in general by demonstrating the ability of SVM-based MVPA to classify individuals with the disorder using the rs-fcMRI. This demonstration was accomplished even after our most conservative motion correction procedures. The work provides several important outcomes, and suggests that there is an ability to classify individuals based on disease status using information available from a brief MRI brain scan. While this particular effort did not focus on maximizing our supervised learning algorithm and feature selection, we were able to make significant disease status predictions well above chance.

We anticipate classification rates with the current ADHD-200 sample will continue to improve with optimization of supervised and unsupervised learning approaches. Optimization and extension of feature sets to include other potential markers is also likely to enhance classification. However, it is only through the future creation of a large-scale datasets, with coordinated recruitment, deep phenotyping, multimodal data acquisition (e.g., rs-fcMRI, diffusion imaging, ASL), and likely improved homogeneity in our subgrouping (Fair et al., 2012) that a fair assessment of the predictive potential of MR-based approaches will be realized. Inclusion of complementary measures from non-MR modalities such as EEG and genetics may further enhance the completeness and accuracy of predictive models.

LIMITATIONS

The findings of the present work should be considered in light of several cautions. Notably, while our movement correction procedures were carefully considered, there is still room for improvement in controlling for this confound. In addition, the present work made use of data aggregated across imaging sites that were not coordinated with respect to their recruitment, diagnostic, or imaging protocols—as such, marked site-related variation exists. Our findings suggest that despite potential differences across centers, there are brain features related to the presence of ADHD and ADHD subtype that are robust to this variation. However, future studies with tightly coordinated imaging and phenotypic data acquisition will be required to replicate the findings of present work and identify additional ADHD-related features that may have been overlooked in the present work due to site-related variation. We note that ADHD relies on strict symptom counts and the DSM-IV system for identifying ADHD-I includes participants who are subthreshold for hyperactivity. Thus, future work may consider examining the extremes (e.g., <3 hyperactive symptoms for ADHD-I) in order to avoid confounds with regard to the imprecise nature of the diagnostic system. Last, we mention that this work does not consider the possibility that the differences noted here could be unique to the effort required to remain still in the ADHD population. Nonetheless, the ability to make valid predictions in individuals with atypical functional neuroanatomy using rs-fcMRI data acquired across multiple institutions provides evidence that this approach can be fruitfully applied in translational studies of disorders with developmental origins.

CONCLUSIONS

Taken together, these results suggest that ADHD-I has distributed atypical connectivity, with prominent findings in control systems, which may underlie the prominent inattentive symptoms in this

population. ADHD-C also shows distributed atypical connectivity, with prominent findings in systems such as the default network, which may mark for an inability to exert top-down attentional control—perhaps in the context of an inability to integrate contextual information supported by default mode processing. Both populations showed atypical nodes in primary sensorimotor systems, supporting previous work implicating this system in ADHD (Mostofsky et al., 2006). These findings point to potential distinct connectivity patterns underlying ADHD subtypes, but also emphasizes the vast heterogeneity in these populations, which will need to be considered in future clinical investigations of ADHD.

ACKNOWLEDGMENTS

We thank Bria L. Thurlow, Corinne A. Stevens, Samuel D. Carpenter, and Yu-Feng Zang for helpful discussions and data preparation. We also extend thanks to the Neuroimaging Informatics Tools and Resources Clearinghouse (http://www.NITRC.org), which hosts the ADHD-200 Consortium's distribution site (http://fcon1000.projects.nitrc.org/indi/adhd200). Last, we acknowledge the multiple sources of funding which supported this work: K99/R00 MH091238 (Fair), R01 MH096773 (Fair), Oregon Clinical and Translational Research Institute (Fair), Medical Research Foundation (Fair), UNCF/Merck Fellowship (Fair), Ford Foundation (Fair), R01 MH086654 (Nigg), R01 HD057076 (Schlaggar), K02 NS053425 (Schlaggar), A. Ernest and Jane G. Stein endowment (Schlaggar), NWO 175.010.2007.010 (Buitelaar), UMC St Radboud (Buitelaar), UMC Groningen en Vrije Universiteit Amsterdam (Buitelaar), R01MH078160 (Mostofsky), R01MH085328 (Mostofsky), Intellectual and Developmental Disabilities Research Center HD-24061 (Mostofsky), P41 RR15241 (Mostofsky), Johns Hopkins CTSA UL1-RR025005 (Mostofsky), R01 MH083320 (Kennedy), P30HD004147 (Kennedy), K01MH082123 (Velanova), R01MH067924 (Luna), R01MH080243 (Luna), Children's Miracle Network—90DD0596/01 (Schweitzer), DHHS Administration for Children and Families (Schweitzer); R03 DA031257 (Schweitzer); R01 MH091068 (Schweitzer); R01MH083246 (Castellanos), R01MH081218 (Castellanos), K23MH087770 (Di Martino), Stavros S. Niarchos Foundation (Castellanos, Milham), Autism Speaks (Castellanos), Leon Levy Foundation (Milham, DiMartino), Phyllis Green and Randolph Cowen endowment (Castellanos), 3K22MH074945-02S1 (Dickstein), Bradley Hospital (Dickstein).

REFERENCES

ADHD-200-Consortium. (2012). The ADHD-200 Consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. *Front. Syst. Neurosci*. 6:62. doi: 10.3389/fnsys.2012.00062

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res*. 125, 279–284.

Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541.

Biswal, B. B., Mennes, M., Zuo, X. N., Gohel, S., Kelly, C., Smith, S. M., et al. (2010). Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4734–4739.

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann. N.Y. Acad. Sci.* 1124, 1–38.

Buckner, R. L., and Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57.

Carp, J. (2011). Optimizing the order of operations for movement scrubbing: comment on Power et al. *Neuroimage.* doi: 10.1016/j.neuroimage.2011.12.061. [Epub ahead of print].

Castellanos, F. X., Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Ghaffari, M., Kirsch, A., et al. (2008). Cingulate-precuneus interactions: a new locus of dysfunction in adult attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 63, 332–337.

Castellanos, F. X., Sonuga-Barke, E. J., Milham, M. P., and Tannock, R. (2006). Characterizing cognition in ADHD: beyond executive dysfunction. *Trends Cogn. Sci.* 10, 117–123.

Castellanos, F. X., Sonuga-Barke, E. J. S., Scheres, A., Di Martino, A., Hyde, C., and Walters, J. R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biol. Psychiatry* 57, 1416–1423.

Costa Dias, T. G., Wilson, V. B., Bathula, D., Iyer, S. P., Mills, K. L., Thurlow, B. L., et al. (2012). Reward circuit connectivity relates to delay discounting in children with attention-deficit/hyperactivity disorder. *Eur. Neuropsychopharmacol.* doi: 10.1016/j.euroneuro.2012.10.015. [Epub ahead of print].

Damoiseaux, J. S., Rombouts, S. A., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., et al. (2006). Consistent resting-state networks across healthy subjects. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13848–13853.

De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., and Formisano, E. (2008). Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *Neuroimage* 43, 44–58.

Denckla, M. B., and Rudel, R. G. (1978). Anomalies of motor development in hyperactive boys. *Ann. Neurol.* 3, 231–233.

Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L., and Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends Cogn. Sci.* 12, 99–105.

Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A. T., et al. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11073–11078.

Dosenbach, N. U., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., et al. (2010). Prediction of individual brain maturity using fMRI. *Science* 329, 1358–1361.

Epstein, J. N., Langberg, J. M., Rosen, P. J., Graham, A., Narad, M. E., Antonini, T. N., et al. (2011). Evidence for higher reaction time variability for children with ADHD on a range of cognitive tasks including reward and event rate manipulations. *Neuropsychology* 25, 427–441.

Fair, D. A., Bathula, D., Nikolas, M. A., and Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6769–6774.

Fair, D. A., Cohen, A. L., Dosenbach, N. U., Church, J. A., Miezin, F. M., Barch, D. M., et al. (2008). The maturing architecture of the brain's default network. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4028–4032.

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U., Church, J. A., Miezin, F. M., et al. (2009). Functional brain networks develop from a "local to distributed" organization. *PLoS Comput. Biol.* 5:e1000381. doi: 10.1371/journal.pcbi.1000381

Fair, D. A., Dosenbach, N. U. F., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., et al. (2007a). Development of distinct control networks through segregation and integration. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13507–13512.

Fair, D. A., Schlaggar, B. L., Cohen, A. L., Miezin, F. M., Dosenbach, N. U., Wenger, K. K., et al. (2007b). A method for using blocked and event-related fMRI data to study "resting state" functional connectivity. *Neuroimage* 35, 396–405.

Fair, D. A., Posner, J., Nagel, B. J., Bathula, D., Dias, T. G., Mills, K. L., et al. (2010). Atypical default network connectivity in youth with attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 68, 1084–1091.

Fair, D. A., and Schlaggar, B. L. (2008). "Brain development," in *Encyclopedia of Brain Development*, eds M. M. Haith and J. B. Benson (San Diego, CA).

Fassbender, C., Zhang, H., Buzy, W. M., Cortes, C. R., Mizuiri, D., Beckett, L., et al. (2009). A lack of default network suppression is linked to increased distractibility in ADHD. *Brain Res.* 1273, 114–128.

Fox, M. D., and Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8, 700–711.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–9678.

Fukunaga, M., Horovitz, S. G., de Zwart, J. A., van Gelderen, P., Balkin, T. J., Braun, A. R., et al. (2008). Metabolic origin of BOLD signal fluctuations in the absence of stimuli. *J. Cereb. Blood Flow. Metab.* 28, 1377–1387.

Fukunaga, M., Horovitz, S. G., van Gelderen, P., de Zwart, J. A., Jansma, J. M., Ikonomidou, V. N., et al. (2006). Large-amplitude, spatially correlated fluctuations in BOLD fMRI signals during extended rest and early sleep stages. *Magn. Reson. Imaging* 24, 979–992.

Gilbert, D. L., Ridel, K. R., Sallee, F. R., Zhang, J., Lipps, T. D., and Wassermann, E. M. (2006). Comparison of the inhibitory and excitatory effects of ADHD medications methylphenidate and atomoxetine on motor cortex. *Neuropsychopharmacology* 31, 442–449.

Greicius, M. (2008). Resting-state functional connectivity in neuropsychiatric disorders. *Curr. Opin. Neurol.* 21, 424–430.

Hagmann, P., Sporns, O., Madan, N., Cammoun, L., Pienaar, R., Wedeen, V. J., et al. (2010). White matter maturation reshapes structural connectivity in the late developing human brain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19067–19072.

Horovitz, S. G., Fukunaga, M., de Zwart, J. A., van Gelderen, P., Fulton, S. C., Balkin, T. J., et al. (2008). Low frequency BOLD fluctuations during resting wakefulness and light sleep: a simultaneous EEG-fMRI study. *Hum. Brain Mapp.* 29, 671–682.

Izawa, J., Pekny, S. E., Marko, M. K., Haswell, C. C., Shadmehr, R., and Mostofsky, S. H. (2012). Motor learning relies on integrated sensory inputs in ADHD, but over-selectively on proprioception in autism spectrum conditions. *Autism Res.* 5, 124–136.

Karrer, B., Levina, E., and Newman, M. E. (2008). Robustness of community structure in networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 77(4 Pt 2):046119. doi: 10.1103/PhysRevE.77.046119

Kelly, A. M., Di Martino, A., Uddin, L. Q., Shehzad, Z., Gee, D. G., Reiss, P. T., et al. (2009). Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cereb. Cortex* 19, 640–657.

Lahey, B. B., and Willcutt, E. G. (2010). Predictive validity of a continuous alternative to nominal subtypes of attention-deficit/hyperactivity disorder for DSM-V. *J. Clin. Child Adolesc. Psychol.* 39, 761–775.

Lancaster, J. L., Glass, T. G., Lankipalli, B. R., Downs, H., Mayberg, H., and Fox, P. T. (1995). A modality-independent approach to spatial normalization of tomographic images of the human brain. *Hum. Brain Mapp.* 3, 209–223.

Leopold, D. A., Murayama, Y., and Logothetis, N. K. (2003). Very slow activity fluctuations in monkey visual cortex: implications for functional brain imaging. *Cereb. Cortex* 13, 422–433.

Matthews, P. M., Honey, G. D., and Bullmore, E. T. (2006). Applications of fMRI in translational medicine and clinical practice. *Nat. Rev. Neurosci.* 7, 732–744.

Meunier, D., Achard, S., Morcom, A., and Bullmore, E. (2009). Age-related changes in modular organization of human brain functional networks. *Neuroimage* 44, 715–723.

Miezin, F., Maccotta, L., Ollinger, J., Petersen, S., and Buckner, R. (2000). Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage* 11, 735–759.

Mills, K. L., Bathula, D., Dias, T. G., Iyer, S. P., Fenesy, M. C., Musser, E. D., et al. (2012). Altered cortico-striatal-thalamic connectivity in relation to spatial working memory capacity in children with ADHD. *Front. Psychiatry* 3:2. doi: 10.3389/fpsyt.2012.00002

Mostofsky, S. H., Newschaffer, C. J., and Denckla, M. B. (2003). Overflow movements predict impaired response inhibition in children with ADHD. *Percept. Mot. Skills* 97(3 Pt 2), 1315–1331.

Mostofsky, S. H., Rimrodt, S. L., Schafer, J. G., Boyce, A., Goldberg, M. C., Pekar, J. J., et al. (2006). Atypical motor and sensory cortex activation in attention-deficit/hyperactivity disorder: a functional magnetic resonance imaging study of simple sequential finger tapping. *Biol. Psychiatry* 59, 48–56.

Mostofsky, S. H., and Simmonds, D. J. (2008). Response inhibition and response selection: two sides of the same coin. *J. Cogn. Neurosci.* 20, 751–761.

Musser, E. D., Backs, R. W., Schmitt, C. F., Ablow, J. C., Measelle, J. R., and Nigg, J. T. (2011). Emotion regulation via the autonomic nervous system in children with attention-deficit/hyperactivity disorder (ADHD). *J. Abnorm. Child Psychol.* 39, 841–852.

Newman, M. E. (2006). Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8577–8582.

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

Nir, Y., Hasson, U., Levy, I., Yeshurun, Y., and Malach, R. (2006). Widespread functional connectivity and fMRI fluctuations in human visual cortex in the absence of visual stimulation. *Neuroimage* 30, 1313–1324.

Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 424–430.

Pennington, B. F., and Ozonoff, S. (1996). Executive functions and developmental psychopathology. *J. Child Psychol. Psychiatry* 37, 51–87.

Pereira, F., Mitchell, T., and Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45(1 Suppl.), S199–S209.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012a). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012b). Steps toward optimizing motion artifact removal in functional connectivity MRI; a reply to Carp. *Neuroimage.* doi: 10.1016/j.neuroimage.2012.03.017. [Epub ahead of print].

Rizk-Jackson, A., Stoffers, D., Sheldon, S., Kuperman, J., Dale, A., Goldstein, J., et al. (2011). Evaluating imaging biomarkers for neurodegeneration in pre-symptomatic Huntington's disease using machine learning techniques. *Neuroimage* 56, 788–796.

Rubinov, M., and Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. *Neuroimage* 56, 2068–2079.

Sagvolden, T., Johansen, E. B., Aase, H., and Russell, V. A. (2005). A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes. *Behav. Brain Sci.* 28, 397–419. discussion: 419–468.

Satterthwaite, T. D., Wolf, D. H., Loughead, J., Ruparel, K., Elliott, M. A., Hakonarson, H., et al. (2012). Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. *Neuroimage* 60, 623–632.

Scholvinck, M. L., Maier, A., Ye, F. Q., Duyn, J. H., and Leopold, D. A. (2010). Neural basis of global resting-state fMRI activity. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10238–10243.

Shannon, B. J., Raichle, M. E., Snyder, A. Z., Fair, D. A., Mills, K. L., Zhang, D., et al. (2011). Premotor functional connectivity predicts impulsivity in juvenile offenders. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11241–11245.

Shaw, P., Eckstrand, K., Sharp, W., Blumenthal, J., Lerch, J. P., Greenstein, D., et al. (2007). Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19649–19654.

Shehzad, Z., Kelly, A. M., Reiss, P. T., Gee, D. G., Gotimer, K., Uddin, L. Q., et al. (2009). The resting brain: unconstrained yet reliable. *Cereb. Cortex* 19, 2209–2229.

Smyser, C. D., Inder, T. E., Shimony, J. S., Hill, J. E., Degnan, A. J., Snyder, A. Z., et al. (2010). Longitudinal analysis of neural network development in preterm infants. *Cereb. Cortex* 20, 2852–2862.

Supekar, K., Musen, M., and Menon, V. (2009). Development of large-scale functional brain networks in children. *PLoS Biol.* 7:e1000157. doi: 10.1371/journal.pbio.1000157

Talairach, J., and Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain.* New York, NY: Thieme Medical Publishers, Inc.

van de Ven, V. G., Formisano, E., Prvulovic, D., Roeder, C. H., and

Linden, D. E. (2004). Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. *Hum. Brain Mapp*. 22, 165–178.

Van Dijk, K. R., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., and Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J. Neurophysiol.* 103, 297–321.

Van Dijk, K. R., Sabuncu, M. R., and Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* 59, 431–438.

Van Essen, D. C. (2005). A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *Neuroimage* 28, 635–662.

Van Essen, D. C., Dickson, J., Harwell, J., Hanlon, D., Anderson, C. H., and Drury, H. A. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *J. Am. Med. Inform. Assoc.* 41, 1359–1378. Available online at: http://brainmap.wustl.edu/caret

Vaurio, R. G., Simmonds, D. J., and Mostofsky, S. H. (2009). Increased intra-individual reaction time variability in attention-deficit/hyperactivity disorder across response inhibition tasks with different cognitive demands. *Neuropsycholgia* 47, 2389–2396.

Vincent, J. L., Patel, G. H., Fox, M. D., Snyder, A. Z., Baker, J. T., Van Essen, D. C., et al. (2007). Intrinsic functional architecture in the anesthetized monkey brain. *Nature* 447, 46–47.

Willcutt, E. G., Nigg, J. T., Pennington, B. F., Solanto, M. V., Rohde, L. A., Tannock, R., et al. (2012). Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. *J. Abnorm. Psychol.* 121, 991–1010.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 July 2012; paper pending published: 24 September 2012; accepted: 30 December 2012; published online: 04 February 2013.

*Citation: Fair DA, Nigg JT, Iyer S, Bathula D, Mills KL, Dosenbach NUF, Schlaggar BL, Mennes M, Gutman D, Bangaru S, Buitelaar JK, Dickstein DP, Di Martino A, Kennedy DN, Kelly C, Luna B, Schweitzer JB, Velanova K, Wang Y-F, Mostofsky S, Castellanos FX and Milham MP (2013) Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data. Front. Syst. Neurosci. **6**:80. doi: 10.3389/fnsys.2012.00080*

Copyright © 2013 Fair, Nigg, Iyer, Bathula, Mills, Dosenbach, Schlaggar, Mennes, Gutman, Bangaru, Buitelaar, Dickstein, Di Martino, Kennedy, Kelly, Luna, Schweitzer, Velanova, Wang, Mostofsky, Castellanos and Milham. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

APPENDIX

PARTICIPANTS AND MEASURES

Data from Brown University, Beijing Normal University, Kennedy Krieger Institute, NYU Child Study Center, and Oregon Health and Science University were aggregated for youth aged 7–14 years. The resulting dataset comprised of 455 control subjects and 193 subjects with a DSM-IV-TR diagnosis of ADHD. A summary of the primary demographics for each sample is provided in **Table 1**. Informed written consent or assent was obtained for all participants, and procedures complied with the Human Investigation Review Board at respective universities. As data for this investigation were aggregated retrospectively (as a large collaborative effort), slightly different ADHD assessment protocols were used across institutions. These procedures are detailed below.

Oregon Health and Science University

Psychiatric diagnoses were based on evaluations with the *Kiddie Schedule for Affective Disorders and Schizophrenia*—KSADS-I (Puig-Antich and Ryan, 1986, Kaufman et al., 1997) administered to a parent; parent and teacher (short form) *Conners' Rating Scale-3rd Edition* (Conners, 2008); and a clinical review by a child psychiatrist and neuropsychologist who had to agree on the diagnosis. Estimates of intelligence were evaluated with a three-subtest short form (Block Design, Vocabulary, and Information) of the *Wechsler Intelligence Scale for Children, Fourth Edition* (WISC-IV) (Wechsler, 1999, 2003).

Children (7–11 years) were excluded if they did not meet criteria for ADHD or non-ADHD groups (i.e., children deemed sub-threshold by the clinicians were excluded). Children were also excluded for an IQ < 80, if a history of neurological illness, chronic medical problems, sensorimotor handicap, autistic disorder, mental retardation, or significant head trauma (with loss of consciousness) was identified by parent report, or if they had evidence of psychotic disorder or bipolar disorder on the structured parent psychiatric interview. Children prescribed short-acting stimulant medications were scanned after a minimum washout of five half-lives (i.e., 24–48 h depending on the preparation). Typically developing control children (TDC) were excluded for presence of conduct disorder, major depressive disorder, or history of psychotic disorder, as well as for presence of ADHD. All children were right handed.

Kennedy Krieger Institute

Psychiatric diagnoses were based on evaluations with the Diagnostic Interview for Children and Adolescents, Fourth Edition [DICA-IV(Reich et al., 1997)], a structured parent interview based on DSM-IV criteria. Estimates of intelligence were evaluated with the WISC-IV (Wechsler, 1999, 2003) and academic achievement was assessed with the Wechsler Individual Achievement Test-II.

All study participants were between 8.0 and 13.0 years, and had a Full Scale IQ of 80 or higher. Children with a specific language disorder or a Reading Disability (RD) were either screened out before a visit or based on school assessment completed within 1 year of participation. RD was based on a statistically significant discrepancy between a child's FSIQ score and his/her Word Reading subtest score from the Wechsler Individual Achievement Test-IIa, or a standard score below 85 on the Word Reading subtest, regardless of IQ score. Participants with visual or hearing impairment, or history of other neurological or psychiatric disorder were excluded.

Children assigned to the ADHD group met criteria for ADHD on the DICA-IV and either had: (1) a T-score of 65 or greater on Conners' Parent Rating Scale-Revised, Long Form [CPRS-R; (Epstein et al., 1997)] subscales "L" (DSM-IV Inattentive) and/or "M" (DSM-IV Hyperactive/Impulsive), or (2) met criteria on the DuPaul ADHD Rating Scale IV (DuPaul et al., 1998) with six out of nine items scored 2 or higher from Inattention items and/or six out of nine scored 2 or higher from the Hyperactivity/Impulsivity items. Children with comorbid DSM-IV diagnoses other than Oppositional Defiant Disorder or Specific Phobias were excluded. DSM-IV criteria and the aforementioned rating scales were also used to evaluate the three ADHD subtypes (Inattentive: ADHD-I; Hyperactive/Impulsive: ADHD-HI; Combined: ADHD-C). Children with ADHD were assigned to the ADHD-I group if they met criteria for inattentiveness but not hyperactivity/impulsivity on the DICA-IV, and had a T-score of 65 or greater on the CPRS Scale L, and a T-score of 60 or less on the CPRS Scale, or had a rating of 2 or 3 on six out of nine Inattention items on the ADHD Rating Scale IV and a rating of 2 or 3 on four or fewer items on the Hyperactivity/Impulsivity scale. Children were assigned to the ADHD-HI if they met criteria for hyperactivity/impulsivity but not inattention on the DICA-IV, and a T-score of 65 or greater on the CPRS Scale M and a T-score of 60 or less on the CPRS Scale L, or had a rating of 2 or 3 on six out of nine Hyperactivity/Impulsivity items on the ADHD Rating Scale IV and a rating of 2 or 3 on four or fewer items on the Inattention scale. All other children who met criteria for ADHD were assigned to the ADHD-C (Combined subtype) group. Children with ADHD taking psychoactive medications other than stimulants were excluded. Children who were taking stimulant medication were asked to withdraw from these medications the day before and the day of testing/scanning.

TDC participants were required to have T-scores of 60 or below on the DSM-IV Inattention (L) and DSM-IV Hyperactivity (M) subscales of CPRS-R and no history of behavioral, emotional, or serious medical problems. Additionally, TDC individuals were not included if there was a history of school-based intervention services as established by parent interview, or if they met DSM-IV psychiatric disorder except specific phobia as reported on the DICA-IV.

New York University

Psychiatric diagnoses were based on evaluations with the Schedule of Affective Disorders and Schizophrenia for Children—Present and Lifetime Version (KSADS-PL) administered to parents and children and the CPRS-LV (Epstein et al., 1997). Intelligence was evaluated with the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999, 2003). Inclusion in the ADHD group required a diagnosis of ADHD based on parent and child responses to the KSADS-PL as well as on a T-score greater than or equal to 60 on at least one ADHD related index of the CPRS-R: LV. ADHD subtype identification was based on interview and review of available records. Psychiatric comorbidity

Table A1 | Node strength values represented in Figures 7 and 8.

Rank	Coordinates X	Y	Z	Procedure 5 TDC vs ADHD-C node strength
1	-35	14	5	28.06
2	-11	-59	15	26.87
3	27	-39	-13	26.39
4	5	-75	-12	21.67
5	-6	44	4	18.37
6	-7	12	36	16.63
7	8	45	21	16.46
8	31	-42	63	14.83
9	-2	23	42	13.79
10	38	-27	54	13.63
11	-26	-43	-11	13.04
12	-1	45	36	10.98
13	-41	42	6	10.73
14	-32	-60	-6	10.60
15	37	36	21	10.46
16	-34	-48	45	10.45
17	5	57	9	10.45
18	10	-26	3	10.25
19	-6	-57	27	10.22
20	41	-15	39	10.10
21	28	-81	12	9.96
22	-29	-30	9	9.89
23	8	0	51	9.86
24	40	-48	21	9.78
25	26	-90	0	7.63
26	-11	39	21	7.51
27	-1	-5	52	7.47
28	51	-47	42	7.43
29	-44	6	15	7.29
30	55	6	18	7.22
31	-53	-45	27	7.14
32	27	50	23	7.10
33	41	-45	9	7.07
34	14	-75	-21	7.05
35	-1	10	46	7.02
36	-10	-72	-15	6.98
37	-47	-48	45	6.84
38	25	42	30	6.84
39	-42	-63	-8	6.84
40	10	-15	8	6.74
41	-46	-63	31	6.74
42	37	-72	12	6.73
43	13	2	10	6.69
44	50	-7	34	6.67
45	56	3	36	6.58
46	-46	-21	48	6.56
47	56	-17	11	6.53
48	4	-52	31	6.50
49	32	-73	-29	6.36
50	48	-61	33	6.30
51	42	-73	27	3.96
52	-22	-54	-21	3.94
53	55	-43	20	3.81
54	-53	3	24	3.79
55	-34	-54	-36	3.78
56	-5	-80	7	3.74
57	-41	3	36	3.73
58	-46	3	3	3.73
59	-25	45	30	3.68
60	-52	-24	9	3.67
61	-51	-38	12	3.57
62	-12	-6	14	3.57
63	-5	-45	24	3.52
64	0	-29	30	3.50
65	-35	-15	15	3.47
66	35	-61	-8	3.46
67	36	16	4	3.45
68	43	-24	45	3.44
69	9	-56	16	3.44
70	44	33	-9	3.44

Rank	Coordinates X	Y	Z	Procedure 5 TDC vs ADHD-I node strength
1	-5	-78	-33	25.58
2	-41	3	36	24.71
3	-43	22	34	21.63
4	41	3	36	17.77
5	2	-66	-24	17.61
6	10	-26	3	14.91
7	-17	23	55	14.87
8	27	50	23	14.78
9	41	-54	45	13.69
10	14	-77	29	13.63
11	37	30	33	13.51
12	35	-61	-8	13.50
13	-9	-72	37	13.12
14	10	-69	39	13.09
15	25	42	30	12.35
16	-17	-51	0	12.03
17	31	-42	63	11.45
18	27	-39	-13	11.10
19	32	-81	-3	10.96
20	27	-74	27	10.87
21	13	2	10	10.80
22	-5	-45	24	10.77
23	-34	-54	-36	10.58
24	-46	-21	48	10.55
25	26	-90	0	10.42
26	-28	51	15	10.30
27	-12	-15	7	10.22
28	37	12	42	10.20
29	-53	-24	36	10.10
30	-32	-60	-6	9.85
31	30	-15	4	7.59
32	-28	-87	5	7.31
33	-6	-57	27	7.29
34	-53	-45	27	7.07
35	8	-45	24	6.92
36	-4	-33	-3	6.73
37	-26	-45	-24	6.68
38	44	33	-9	6.63
39	-4	-40	43	6.61
40	-23	-60	-33	6.56
41	52	-33	-15	6.51
42	40	-48	21	6.51
43	55	-43	20	6.43
44	-44	6	15	6.34
45	-8	-42	3	4.31
46	-25	45	30	4.28
47	56	-17	11	4.03
48	10	-15	8	3.99
49	-29	-16	2	3.97
50	-3	25	30	3.85
51	5	-75	-12	3.80
52	-40	-42	39	3.80
53	0	-29	30	3.80
54	55	-7	20	3.77
55	-11	-59	15	3.77
56	-35	-82	-5	3.77
57	8	-76	12	3.75
58	55	6	18	3.73
59	7	-42	48	3.70
60	-42	-63	-8	3.68
61	8	0	51	3.68
62	-41	42	6	3.65
63	41	-15	39	3.65
64	43	-12	26	3.65
65	-35	14	5	3.59
66	50	-18	-9	3.48
67	40	-27	18	3.47
68	-26	-12	53	3.47
69	-38	-18	57	3.45
70	-46	3	3	3.42

Rank	Coordinates X	Y	Z	Procedure 5 ADHD-I vs ADHD-C node strength
1	37	-72	12	14.00
2	5	-75	-12	13.15
3	-26	-12	53	11.55
4	-7	12	36	10.69
5	-26	-43	-11	10.44
6	5	57	9	10.11
7	27	-39	-13	9.04
8	7	37	0	8.40
9	44	33	-9	7.80
10	-4	-33	-3	7.61
11	-32	-60	-6	7.57
12	14	-75	-21	7.41
13	27	-74	27	7.17
14	37	36	21	6.81
15	8	45	21	5.94
16	32	-81	-3	5.81
17	-35	14	5	5.70
18	-42	-63	-8	5.62
19	4	-52	31	5.59
20	-1	-5	52	5.47
21	27	50	23	5.09
22	36	16	4	4.90
23	2	-66	-24	4.90
24	25	42	30	4.55
25	40	-27	18	4.46
26	49	-33	7	4.30
27	-43	22	34	4.18
28	-32	-66	-29	4.18
29	-22	-54	-21	3.94
30	-35	-69	36	3.90
31	55	-7	20	3.85
32	55	-43	20	3.81
33	41	-45	9	3.81
34	48	-61	33	3.74
35	-41	42	6	3.61
36	18	-66	-2	3.57
37	-46	-21	48	3.57
38	13	-90	0	3.51
39	35	-61	-8	3.51
40	-29	-16	2	3.50
41	-11	-59	15	3.36
42	-29	-30	9	3.22
43	-3	25	30	3.22
44	-5	-78	-33	3.19
45	-35	-82	-5	3.09
46	-46	3	3	3.07
47	8	33	24	2.91
48	8	-45	24	2.84
49	-12	-15	7	2.78
50	41	-15	39	2.65
51	28	-81	12	2.63
52	30	-15	4	2.57
53	-10	-72	-15	2.56
54	-38	-18	57	2.55
55	19	-78	-3	2.54
56	32	27	12	2.44
57	37	12	42	2.37
58	-51	-38	12	2.31
59	-19	-78	-33	2.31
60	-4	-93	9	2.28
61	14	-77	29	2.12
62	21	27	49	2.10
63	-15	-64	-21	2.06
64	-28	51	15	2.02
65	-6	-57	27	1.89
66	-17	-51	0	1.88
67	-46	-63	31	1.74
68	18	-80	-33	1.68
69	-40	-76	22	1.64
70	-44	6	15	1.57

(Continued)

Table A1 | Continued

Rank	Coordinates X	Y	Z	Procedure 5 TDC vs ADHD-C node strength
71	41	3	36	3.42
72	7	37	0	3.41
73	-40	-6	12	3.40
74	18	-66	-2	3.37
75	-40	-33	45	3.36
76	10	-69	39	3.36
77	-26	-12	53	3.36
78	-32	-66	-29	3.31
79	-17	23	55	3.29
80	16	-69	18	3.28
81	55	-7	20	3.27
82	31	-15	18	3.26
83	31	-61	-29	3.25
84	-26	-45	-24	3.25
85	-3	25	30	3.25
86	-17	-51	0	3.25
87	-46	-15	35	3.23
88	8	-76	12	3.22
89	32	-81	-3	3.21
90	14	-77	29	3.20
91	-28	-75	24	3.20
92	-4	-33	-3	3.18
93	-8	-42	3	3.17
94	-52	-12	23	3.17
95	49	-33	7	3.15

Rank	Coordinates X	Y	Z	Procedure 5 TDC vs ADHD-I node strength
71	-5	-80	7	3.42
72	9	-56	16	3.39
73	40	42	3	3.36
74	51	-47	42	3.35
75	-6	44	4	3.33
76	-10	-72	-15	3.32
77	36	16	4	3.27
78	-52	-24	9	3.27
79	32	-73	-29	3.25
80	-32	-66	-29	3.25
81	-52	-12	23	3.24
82	-40	-76	22	3.23
83	14	-75	-21	3.19
84	-35	-15	15	3.16
85	4	-52	31	3.16
86	-35	-69	36	3.16
87	-26	-43	-11	3.15
88	-50	23	19	3.14
89	7	37	0	3.14
90	31	-61	-29	3.14
91	-1	45	36	3.14
92	-58	-42	-3	3.14
93	-40	-6	12	3.14
94	-5	-53	16	3.13
95	-53	3	24	3.13
96	-46	-15	35	3.12
97	13	-90	0	3.12

Rank	Coordinates X	Y	Z	Procedure 5 ADHD-I vs ADHD-C node strength
71	-25	45	30	1.52
72	-28	-75	24	1.49
73	31	-15	18	1.47
74	-1	45	36	1.44
75	-17	23	55	1.37
76	-26	-45	-24	1.36
77	-53	-45	27	1.35
78	-53	3	24	1.35
79	-9	-72	37	1.33
80	-2	-75	29	1.29
81	-12	-6	14	1.22
82	-11	39	21	1.19
83	13	2	10	1.11
84	-39	-38	15	1.01
85	-8	-42	3	1.01
86	-43	-9	48	1.01
87	-34	-48	45	0.85
88	-31	-59	42	0.74
89	10	-15	8	0.73
90	38	-27	54	0.70
91	-56	-26	-14	0.64
92	40	42	3	0.63
93	42	-73	27	0.54
94	-34	-54	-36	0.50
95	32	-73	-29	0.50
96	10	-69	39	0.48
97	-41	3	36	0.48
98	7	-42	48	0.33
99	37	30	33	0.28

Rank	Coordinates X	Y	Z	Procedure 8 TDC vs ADHD-C node strength
1	27	-39	-13	30.47
2	-26	-43	-11	27.79
3	5	57	9	20.62
4	5	-75	-12	17.95
5	19	-78	-3	17.84
6	-32	-60	-6	16.96
7	-11	-59	15	16.85
8	-41	3	36	16.78
9	-10	-72	-15	14.93
10	40	-48	21	13.07
11	-11	39	21	10.68
12	-53	3	24	10.41
13	2	-66	-24	10.29
14	-6	44	4	10.26
15	-1	45	36	10.16
16	-35	14	5	10.12
17	55	6	18	10.11
18	42	-73	27	9.96
19	-5	-80	7	9.94
20	8	0	51	9.72
21	8	45	21	9.60
22	18	-66	-2	7.74
23	55	-43	20	7.71
24	36	16	4	7.56
25	31	-42	63	7.46
26	-35	-15	15	7.30
27	-41	42	6	7.21
28	-22	-54	-21	7.18
29	-32	-66	-29	7.10
30	-29	-30	9	7.02
31	32	27	12	6.97
32	-6	-57	27	6.95
33	-44	6	15	6.95
34	-12	-6	14	6.80

Rank	Coordinates X	Y	Z	Procedure 8 TDC vs ADHD-I node strength
1	-41	3	36	28.33
2	-5	-78	-33	21.12
3	37	30	33	20.23
4	-46	-21	48	19.93
5	-34	-54	-36	18.40
6	-17	23	55	18.33
7	-23	-60	-33	17.45
8	14	-77	29	15.96
9	41	-54	45	13.87
10	41	3	36	13.84
11	32	-81	-3	13.69
12	27	50	23	13.56
13	-5	-45	24	13.37
14	-10	-72	-15	12.95
15	-43	22	34	12.26
16	2	-66	-24	11.37
17	-3	25	30	11.15
18	-40	-48	27	10.49
19	35	-61	-8	10.39
20	30	-15	4	10.29
21	7	-42	48	9.65
22	-32	-60	-6	9.56
23	-25	45	30	9.48
24	49	-33	7	9.43
25	-17	-51	0	8.68
26	25	42	30	7.69
27	-53	-45	27	7.55
28	-32	-66	-29	7.55
29	-6	-57	27	7.54
30	-12	-15	7	7.45
31	-41	42	6	7.33
32	-5	-80	7	7.25
33	40	-27	18	7.15
34	56	-17	11	7.12

Rank	Coordinates X	Y	Z	Procedure 8 ADHD-I vs ADHD-C node strength
1	-26	-43	-11	28.20
2	27	-39	-13	19.94
3	37	-72	12	17.98
4	5	57	9	14.70
5	32	-81	-3	12.66
6	5	-75	-12	8.86
7	19	-78	-3	8.80
8	4	-52	31	8.12
9	-26	-12	53	8.06
10	-6	-57	27	7.70
11	44	33	-9	7.30
12	41	-15	39	6.94
13	36	16	4	6.64
14	-17	-51	0	6.57
15	-10	-72	-15	6.43
16	-53	3	24	6.10
17	-22	-54	-21	5.79
18	27	-74	27	5.44
19	-32	-60	-6	5.43
20	27	50	23	5.39
21	40	-27	18	5.30
22	-46	-63	31	5.25
23	37	12	42	5.16
24	18	-66	-2	5.09
25	-35	-82	-5	4.97
26	-5	-45	24	4.93
27	-38	-18	57	4.93
28	-43	-9	48	4.45
29	-43	22	34	4.32
30	56	-17	11	4.32
31	-46	-21	48	4.27
32	55	-43	20	4.25
33	41	-45	9	4.25
34	16	-30	60	3.69

(Continued)

Table A1 | Continued

Rank	Coordinates X	Y	Z	Procedure 8 TDC vs ADHD-C node strength
35	44	-63	5	6.68
36	56	-17	11	6.57
37	31	-15	18	6.51
38	13	2	10	6.51
39	-32	-66	-29	6.40
40	37	-72	12	6.39
41	41	-3	15	6.38
42	41	-45	9	4.25
43	41	-15	39	4.18
44	-15	-64	-21	4.06
45	-1	-5	52	4.06
46	-34	-48	45	3.95
47	-5	-60	-15	3.95
48	-43	-9	48	3.86
49	-40	-33	45	3.74
50	30	-61	39	3.74
51	14	-75	-21	3.73
52	38	-27	54	3.69
53	-42	-63	-8	3.65
54	-43	22	34	3.61
55	-4	-33	-3	3.59
56	8	33	24	3.54
57	-40	-6	12	3.51
58	40	-27	18	3.51
59	4	-52	31	3.49
60	-46	3	3	3.48
61	8	15	36	3.48
62	10	-15	8	3.45
63	52	-33	-15	3.42
64	-8	-42	3	3.39
65	-4	-93	9	3.38
66	-2	23	42	3.37
67	10	-26	3	3.36
68	-53	-45	27	3.34
69	-26	-45	-24	3.34
70	0	-29	30	3.34
71	-12	-15	7	3.33
72	-47	-48	45	3.33
73	43	-12	26	3.33
74	-1	10	46	3.32
75	-19	2	9	3.30
76	41	3	36	3.30
77	31	-61	-29	3.30
78	25	42	30	3.27
79	-23	-60	-33	3.27
80	-3	25	30	3.25
81	21	27	49	3.23
82	37	36	21	3.22
83	48	18	12	3.21
84	35	-61	-8	3.19
85	-17	-51	0	3.17
86	32	-73	-29	3.17
87	10	-69	39	3.17
88	-17	23	55	3.16
89	43	22	34	3.15
90	31	-61	-29	3.14
91	-46	-15	35	3.13
92	-26	-12	53	3.13
93	14	-77	29	3.12
94	-50	23	19	3.11
95	32	-81	-3	3.10

Rank	Coordinates X	Y	Z	Procedure 8 TDC vs ADHD-I node strength
35	55	-7	20	6.98
36	-28	-87	5	6.91
37	-8	-42	3	6.86
38	9	-56	16	6.76
39	-35	-82	-5	6.61
40	-1	-5	52	6.53
41	-58	-42	-3	6.52
42	-1	45	36	6.48
43	-40	-42	39	6.44
44	-53	-24	36	6.39
45	28	-81	12	6.32
46	-53	3	24	6.18
47	8	-76	12	4.90
48	26	-90	0	4.33
49	32	-73	-29	4.25
50	37	12	42	3.96
51	-29	-16	2	3.87
52	-26	-12	53	3.86
53	55	6	18	3.77
54	44	33	-9	3.73
55	31	-42	63	3.69
56	27	-74	27	3.61
57	0	-29	30	3.55
58	-6	44	4	3.47
59	-38	-18	57	3.46
60	10	-15	8	3.45
61	32	27	12	3.43
62	40	-48	21	3.41
63	44	-63	5	3.41
64	8	33	24	3.41
65	-46	-15	35	3.39
66	48	-61	33	3.39
67	-4	-33	-3	3.37
68	-56	-26	-14	3.37
69	-9	-72	37	3.33
70	5	-75	-12	3.31
71	-40	-76	22	3.29
72	18	-80	-33	3.29
73	-50	-63	12	3.28
74	-4	-40	43	3.28
75	4	-52	31	3.27
76	-51	-38	12	3.25
77	-11	-59	15	3.24
78	-19	2	9	3.24
79	-46	3	3	3.24
80	10	-26	3	3.23
81	7	37	0	3.21
82	10	-69	39	3.21
83	-28	-75	24	3.17
84	8	0	51	3.17
85	38	-27	54	3.16
86	-29	-30	9	3.16
87	-40	-6	12	3.16
88	13	2	10	3.15
89	30	-61	39	3.15
90	37	36	21	3.14
91	27	-39	-13	3.14
92	19	-78	-3	3.13
93	50	-18	-9	3.12
94	-35	14	5	3.10
95	51	-47	42	3.08
96	-43	-9	48	3.08
97	41	-15	39	3.05
98	8	-45	24	3.04

Rank	Coordinates X	Y	Z	Procedure 8 ADHD-I vs ADHD-C node strength
35	-12	-15	7	3.57
36	8	15	36	3.48
37	13	-90	0	3.41
38	-28	51	15	3.23
39	-11	-59	15	3.13
40	-41	42	6	2.91
41	-8	-42	3	2.90
42	-35	14	5	2.90
43	30	-15	4	2.89
44	-15	-64	-21	2.83
45	-24	-33	61	2.79
46	-40	-33	45	2.79
47	38	-27	54	2.75
48	-19	2	9	2.69
49	32	27	12	2.69
50	56	3	36	2.55
51	55	-7	20	2.53
52	35	-61	-8	2.51
53	-32	-66	-29	2.51
54	16	-69	18	2.42
55	26	-90	0	2.40
56	-16	-76	29	2.30
57	-11	39	21	2.29
58	14	-77	29	2.27
59	18	-80	-33	2.26
60	41	-3	15	2.17
61	-39	-38	15	2.14
62	21	27	49	1.96
63	-7	12	36	1.95
64	-4	-33	-3	1.93
65	-44	6	15	1.90
66	-5	-78	-33	1.84
67	-17	23	55	1.69
68	49	-33	7	1.68
69	-28	-75	24	1.65
70	-56	-26	-14	1.57
71	-5	-53	16	1.46
72	-3	25	30	1.37
73	50	-7	34	1.21
74	50	-18	-9	1.12
75	40	42	3	1.04
76	-58	-42	-3	0.97
77	-4	-93	9	0.96
78	25	42	30	0.93
79	-35	-69	36	0.80
80	-40	-42	39	0.78
81	-34	-54	-36	0.74
82	32	-73	-29	0.74
83	48	-61	33	0.73
84	-1	45	36	0.59
85	43	22	34	0.59
86	-29	-16	2	0.40
87	-53	-24	36	0.21
88	-6	44	4	0.16
89	10	-69	39	0.06
90	-41	3	36	0.06
91	-53	-45	27	0.04

(Continued)

Table A1 | Continued

Rank	Coordinates X	Y	Z	Procedure 10 TDC vs ADHD-C node strength
1	-35	14	5	27.84
2	-11	-59	15	26.73
3	27	-39	-13	26.24
4	5	-75	-12	21.43
5	38	-27	54	16.56
6	8	45	21	16.27
7	-6	44	4	15.05
8	31	-42	63	14.72
9	-2	23	42	13.68
10	5	57	9	13.58
11	-7	12	36	13.09
12	-26	-43	-11	13.00
13	-1	45	36	10.86
14	-32	-60	-6	10.63
15	-41	42	6	10.59
16	37	36	21	10.45
17	-44	6	15	10.35
18	-34	-48	45	10.29
19	-6	-57	27	10.25
20	41	-15	39	10.07
21	10	-26	3	10.00
22	56	-17	11	9.66
23	40	-48	21	9.66
24	8	0	51	9.66
25	26	-90	0	7.64
26	-1	-5	52	7.42
27	-11	39	21	7.41
28	51	-47	42	7.36
29	55	6	18	7.27
30	27	50	23	7.11
31	-53	-45	27	7.06
32	-10	-72	-15	6.95
33	41	-45	9	6.90
34	25	42	30	6.90
35	-1	10	46	6.88
36	-42	-63	-8	6.83
37	-47	-48	45	6.77
38	14	-75	-21	6.73
39	37	-72	12	6.72
40	-46	-63	31	6.71
41	-29	-30	9	6.66
42	28	-81	12	6.65
43	36	16	4	6.64
44	9	-56	16	6.63
45	13	2	10	6.63
46	50	-7	34	6.61
47	41	3	36	6.54
48	10	-15	8	6.53
49	56	3	36	6.51
50	4	-52	31	6.51
51	-4	-33	-3	6.33
52	32	-73	-29	6.32
53	42	-73	27	3.92
54	-22	-54	-21	3.90
55	-53	3	24	3.76
56	-46	3	3	3.74
57	-25	45	30	3.72
58	-5	-80	7	3.71
59	-34	-54	-36	3.70
60	-41	3	36	3.70
61	-52	-24	9	3.70
62	55	-43	20	3.69
63	-12	-6	14	3.54
64	-5	-45	24	3.51
65	-51	-38	12	3.51
66	43	-24	45	3.49
67	-46	-21	48	3.46
68	0	-29	30	3.43
69	-35	-15	15	3.43
70	35	-61	-8	3.42

Rank	Coordinates X	Y	Z	Procedure 10 TDC vs ADHD-I node strength
1	-41	3	36	27.87
2	-5	-78	-33	21.70
3	41	3	36	21.51
4	27	50	23	18.34
5	-43	22	34	18.21
6	-5	-45	24	17.52
7	37	30	33	16.41
8	10	-69	39	15.99
9	31	-42	63	14.96
10	-34	-54	-36	14.11
11	-46	-21	48	14.09
12	14	-77	29	13.81
13	41	-54	45	13.59
14	35	-61	-8	12.86
15	25	42	30	12.20
16	-17	-51	0	11.72
17	-17	23	55	11.42
18	10	-26	3	11.18
19	27	-74	27	10.96
20	27	-39	-13	10.87
21	13	2	10	10.58
22	2	-66	-24	10.53
23	26	-90	0	10.40
24	-28	51	15	10.08
25	-12	-15	7	9.94
26	37	12	42	9.78
27	51	-47	42	9.45
28	-32	-60	-6	9.43
29	-6	-57	27	7.41
30	30	-15	4	7.41
31	-28	-87	5	7.37
32	32	-81	-3	7.34
33	-40	-42	39	7.29
34	-53	-45	27	7.15
35	-53	-24	36	7.07
36	-23	-60	-33	7.06
37	-3	25	30	6.96
38	55	-43	20	6.81
39	8	-45	24	6.81
40	40	-48	21	6.68
41	-4	-33	-3	6.67
42	-32	-66	-29	6.64
43	50	-18	-9	6.41
44	-4	-40	43	6.37
45	9	-56	16	6.34
46	-26	-45	-24	6.33
47	-53	3	24	6.32
48	-8	-42	3	4.43
49	56	-17	11	4.09
50	5	-75	-12	4.03
51	-35	-82	-5	3.99
52	10	-15	8	3.87
53	8	-76	12	3.84
54	-35	14	5	3.78
55	-25	45	30	3.73
56	40	-27	18	3.72
57	-26	-12	53	3.72
58	55	-7	20	3.70
59	-11	-59	15	3.70
60	-41	42	6	3.68
61	-29	-16	2	3.66
62	55	6	18	3.65
63	0	-29	30	3.64
64	-42	-63	-8	3.61
65	8	0	51	3.61
66	44	33	-9	3.51
67	32	-73	-29	3.51
68	7	-42	48	3.48
69	40	42	3	3.39
70	36	16	4	3.39

Rank	Coordinates X	Y	Z	Procedure 10 ADHD-I vs ADHD-C node strength
1	37	-72	12	15.28
2	27	-39	-13	14.24
3	-26	-12	53	10.28
4	5	57	9	7.96
5	5	-75	-12	7.91
6	27	50	23	7.83
7	-4	-33	-3	7.54
8	4	-52	31	7.35
9	-46	3	3	7.28
10	-26	-43	-11	7.10
11	27	-74	27	7.08
12	2	-66	-24	6.30
13	-11	-59	15	5.84
14	32	-81	-3	5.70
15	-35	14	5	5.68
16	8	45	21	5.38
17	37	12	42	5.26
18	13	-90	0	5.25
19	-1	45	36	5.16
20	-6	-57	27	5.06
21	44	33	-9	4.89
22	36	16	4	4.87
23	48	18	12	4.63
24	-7	12	36	4.48
25	40	-27	18	4.38
26	-46	-21	48	4.33
27	-38	-18	57	4.30
28	-53	3	24	4.19
29	-22	-54	-21	3.90
30	-32	-60	-6	3.86
31	-35	-69	36	3.84
32	14	-75	-21	3.81
33	55	-7	20	3.75
34	55	-43	20	3.69
35	41	-45	9	3.69
36	-41	42	6	3.50
37	-29	-16	2	3.50
38	-43	22	34	3.40
39	-44	6	15	3.36
40	-28	-75	24	3.35
41	-1	-5	52	3.22
42	8	33	24	2.86
43	8	-45	24	2.80
44	-5	-78	-33	2.79
45	-12	-15	7	2.77
46	40	42	3	2.76
47	41	-15	39	2.63
48	30	-15	4	2.58
49	-10	-72	-15	2.55
50	28	-81	12	2.52
51	-2	23	42	2.51
52	35	-61	-8	2.48
53	-32	-66	-29	2.48
54	19	-78	-3	2.45
55	49	-33	7	2.29
56	-4	-93	9	2.23
57	-35	-82	-5	2.23
58	56	-17	11	2.21
59	-52	-12	23	2.16
60	41	-3	15	2.16
61	14	-77	29	2.13
62	-15	-64	-21	2.08
63	16	-30	60	2.05
64	-8	-42	3	2.05
65	-28	51	15	1.92
66	-42	-63	-8	1.90
67	25	42	30	1.81
68	-46	-63	31	1.77
69	-40	-76	22	1.61
70	18	-66	-2	1.60

(Continued)

Table A1 | Continued

Rank	Coordinates X	Y	Z	Procedure 10 TDC vs ADHD-C node strength
71	-40	-6	12	3.41
72	44	33	-9	3.39
73	7	37	0	3.38
74	-26	-12	53	3.33
75	-32	-66	-29	3.32
76	55	-7	20	3.32
77	18	-66	-2	3.32
78	16	-69	18	3.30
79	-17	23	55	3.28
80	-40	-33	45	3.28
81	10	-69	39	3.28
82	-17	-51	0	3.23
83	-46	-15	35	3.23
84	31	-15	18	3.23
85	-26	-45	-24	3.22
86	14	-77	29	3.21
87	-28	-75	24	3.21
88	31	-61	-29	3.20
89	49	-33	7	3.18
90	-3	25	30	3.18
91	32	-81	-3	3.15
92	48	-61	33	3.14
93	-52	-12	23	3.13
94	8	-76	12	3.13
95	-56	-26	-14	3.13

Rank	Coordinates X	Y	Z	Procedure 10 TDC vs ADHD-I node strength
71	-5	-80	7	3.38
72	32	27	12	3.38
73	-34	-48	45	3.37
74	-10	-72	-15	3.36
75	-26	-43	-11	3.34
76	31	-61	-29	3.31
77	-40	-76	22	3.28
78	-52	-12	23	3.21
79	52	-33	-15	3.21
80	-35	-15	15	3.21
81	4	-52	31	3.21
82	-46	-15	35	3.21
83	13	-90	0	3.21
84	-5	-53	16	3.20
85	-1	45	36	3.19
86	-58	-42	-3	3.19
87	-46	3	3	3.19
88	37	-72	12	3.18
89	41	-3	15	3.18
90	-52	-24	21	3.15
91	-38	-18	57	3.12
92	38	-27	54	3.12
93	-35	-69	36	3.11
94	-16	-76	29	3.11
95	48	18	12	3.11
96	-9	-72	37	3.08
97	-6	44	4	3.08

Rank	Coordinates X	Y	Z	Procedure 10 ADHD-I vs ADHD-C node strength
71	21	27	49	1.57
72	-40	-42	39	1.56
73	48	-61	33	1.42
74	-53	-45	27	1.36
75	-17	23	55	1.34
76	8	0	51	1.33
77	-2	-75	29	1.27
78	-9	-72	37	1.27
79	-12	-6	14	1.18
80	18	-80	-33	1.18
81	13	2	10	1.15
82	-34	-48	45	0.86
83	10	-15	8	0.83
84	-3	25	30	0.67
85	38	-27	54	0.66
86	-56	-26	-14	0.62
87	50	-18	-9	0.62
88	-34	-54	-36	0.50
89	32	-73	-29	0.50
90	10	-69	39	0.49
91	-41	3	36	0.49
92	55	6	18	0.44
93	7	-42	48	0.33
94	37	30	33	0.30
95	37	36	21	0.14
96	50	-7	34	0.12
97	35	-6	0	0.12

Rank	Coordinates X	Y	Z	Age - Procedure 5 FDR	Concensus
1	30	-15	4	77.18	15.09
2	-29	-16	2	76.76	14.77
3	-19	2	9	71.04	14.58
4	13	2	10	70.05	11.24
5	-6	-57	27	61.32	7.23
6	14	-77	29	60.78	7.58
7	16	-69	18	60.69	6.85
8	-53	-45	27	60.34	8.11
9	-35	14	5	60.23	7.29
10	10	-26	3	59.34	8.96
11	10	-69	39	58.58	6.61
12	-12	-6	14	58.06	8.68
13	-4	-33	-3	57.33	8.67
14	-46	3	3	56.45	7.97
15	-5	-53	16	55.22	6.56
16	10	-15	8	54.36	9.14
17	36	16	4	51.73	5.20
18	4	-52	31	51.69	5.42
19	38	-27	54	51.69	6.94
20	31	-42	63	51.01	6.28
21	-1	45	36	49.49	6.07
22	-7	12	36	49.40	6.09
23	51	-47	42	49.13	6.84
24	-38	-30	57	49.11	5.44
25	55	-43	20	49.10	2.90
26	8	45	21	47.70	7.25
27	8	0	51	47.54	4.74
28	-6	44	4	47.39	6.26
29	-52	-12	23	47.24	8.00
30	-1	-5	52	47.06	4.67
31	-26	-12	53	45.93	3.89
32	-24	-33	61	45.84	3.08
33	41	-3	15	45.60	3.52
34	25	42	30	45.49	2.26
35	-29	-30	9	45.34	3.83
36	-12	-15	7	45.19	4.75
37	-35	-69	36	44.49	5.48
38	-38	-18	57	44.21	5.97

Rank	Coordinates X	Y	Z	Age - Procedure 7 FDR	Concensus
1	-29	-16	2	18.40	15.99
2	30	-15	4	16.48	14.20
3	-52	-12	23	14.14	12.63
4	-53	-45	27	13.16	10.71
5	-19	2	9	12.69	11.35
6	13	2	10	12.65	11.33
7	14	-77	29	12.31	11.53
8	10	-26	3	11.74	11.14
9	-6	-57	27	10.95	9.17
10	-46	3	3	10.23	8.88
11	-35	-15	15	9.75	8.44
12	31	-42	63	9.30	8.00
13	-7	12	36	9.29	7.30
14	51	-47	42	9.25	8.60
15	-1	-5	52	9.08	7.16
16	-35	14	5	8.69	6.71
17	10	-15	8	8.29	7.80
18	8	0	51	8.21	7.00
19	36	16	4	8.09	7.45
20	38	-27	54	7.90	7.39
21	-38	-30	57	7.86	5.90
22	4	-52	31	7.72	6.89
23	-38	-18	57	7.48	6.95
24	-35	-69	36	7.33	6.09
25	48	18	12	7.18	6.03
26	-5	-53	16	7.13	5.76
27	-29	-30	9	6.72	4.94
28	-12	-6	14	6.51	5.46
29	8	45	21	6.49	5.21
30	5	57	9	6.44	5.96
31	35	-6	0	6.41	5.35
32	32	-73	-29	6.39	4.62
33	18	-66	-2	6.28	5.95
34	-17	-51	0	6.19	5.83
35	-46	-15	35	6.01	5.67
36	-1	45	36	5.93	4.79
37	16	-69	18	5.88	4.84
38	10	-69	39	5.88	5.50

Rank	Coordinates X	Y	Z	Age - Procedure 10 FDR	Concensus
1	-19	2	9	68.60	0.00
2	-29	-16	2	59.92	7.00
3	14	-77	29	54.10	8.53
4	-35	14	5	53.11	12.59
5	30	-15	4	50.79	9.03
6	-6	-57	27	49.82	9.22
7	16	-69	18	47.17	0.00
8	51	-47	42	46.66	2.04
9	-26	-12	53	46.58	7.30
10	13	2	10	46.50	3.29
11	38	-27	54	45.52	4.26
12	-53	-45	27	44.93	8.48
13	5	57	9	44.49	6.84
14	-52	-12	23	43.82	0.00
15	10	-69	39	43.75	3.82
16	-12	-6	14	42.65	0.00
17	-38	-30	57	42.52	0.00
18	-5	-53	16	42.51	0.00
19	10	-26	3	42.06	4.10
20	36	16	4	42.04	0.00
21	8	0	51	41.28	0.00
22	-38	-18	57	40.81	0.00
23	31	-42	63	40.33	3.65
24	-7	12	36	39.71	3.22
25	-4	-33	-3	39.10	7.55
26	10	-15	8	38.26	2.76
27	4	-52	31	38.12	7.93
28	-1	-5	52	37.94	1.99
29	-1	45	36	37.74	4.90
30	-46	3	3	37.29	6.86
31	7	-42	48	36.23	3.22
32	-24	-33	61	36.06	0.00
33	-35	-15	15	35.20	0.00
34	48	-61	33	35.02	0.00
35	41	-3	15	34.93	0.00
36	-50	23	19	34.64	0.00
37	35	-6	0	33.76	0.00
38	18	-66	-2	33.35	3.55

(Continued)

Table A1 | Continued

Rank	Coordinates X	Y	Z	Age - Procedure 5 FDR	Concensus
39	35	-6	0	42.87	6.62
40	-35	-15	15	42.59	6.85
41	5	57	9	42.15	5.13
42	7	-42	48	40.76	4.17
43	-46	-63	31	39.80	2.81
44	48	-61	33	39.50	3.05
45	-46	-15	35	39.34	4.54
46	18	-66	-2	39.11	5.72
47	48	18	12	38.21	5.46
48	55	6	18	38.01	5.63
49	-8	-42	3	37.99	2.41
50	-53	-24	36	37.93	2.17
51	-17	23	55	37.14	2.42
52	-34	-54	-36	36.31	3.69
53	43	-12	26	35.76	2.25
54	55	-7	20	34.16	4.16
55	-1	10	46	32.77	3.77
56	21	27	49	32.71	1.45
57	-17	-51	0	31.49	5.63
58	40	-27	18	31.48	4.77
59	-52	-24	9	31.13	2.20
60	-50	23	19	30.88	2.12
61	49	-33	7	30.53	1.66
62	41	3	36	30.52	4.42
63	-31	-59	42	29.78	3.63
64	-46	-21	48	29.45	2.26
65	43	-24	45	28.94	4.38
66	-40	-6	12	28.74	1.61
67	31	-15	18	28.67	4.37
68	-23	-60	-33	28.59	1.44
69	27	-39	-13	27.80	0.78
70	-26	-43	-11	27.56	2.05
71	8	15	36	27.44	1.34
72	-40	-33	45	27.11	1.50
73	-47	-48	45	27.04	0.68
74	8	33	24	26.52	2.17
75	-58	-42	-3	26.51	0.71
76	41	-54	45	26.42	2.25
77	-51	-51	36	26.29	2.10
78	40	-48	21	26.05	1.40
79	-32	-60	-6	25.91	0.71
80	16	-30	60	25.78	2.98
81	-11	-59	15	25.72	2.93
82	-32	-66	-29	25.23	0.77
83	-41	42	6	24.94	0.66
84	18	-80	-33	24.73	2.33
85	-16	-76	29	24.42	1.57
86	37	30	33	23.93	0.65
87	56	-17	11	23.66	3.19
88	-43	22	34	23.44	0.69
89	-11	39	21	23.40	0.00
90	27	-74	27	23.11	0.00
91	8	-45	24	23.10	2.04
92	44	-63	5	22.90	0.00
93	-9	-72	37	22.46	0.77
94	37	36	21	22.23	0.77
95	13	-90	0	22.14	1.44
96	50	-7	34	21.60	1.44
97	35	-61	-8	21.43	1.50
98	41	-15	39	21.38	0.74
99	37	-72	12	21.24	0.76
100	-42	-63	-8	21.19	0.72
101	-34	-48	45	21.13	0.64
102	-43	-9	48	20.79	0.00
103	37	12	42	20.74	0.70
104	-52	-24	21	20.68	2.29
105	-5	-78	-33	20.11	0.73
106	41	-45	9	19.66	0.77
107	-28	-75	24	19.42	0.00
108	44	33	-9	19.39	1.53
109	9	-56	16	19.12	2.88
110	50	-18	-9	19.08	0.00
111	7	37	0	18.98	0.00
112	40	42	3	18.94	1.51

Rank	Coordinates X	Y	Z	Age - Procedure 7 FDR	Concensus
39	43	-24	45	5.83	4.06
40	43	-12	26	5.73	4.72
41	-1	10	46	5.69	4.60
42	41	3	36	5.66	4.58
43	31	-15	18	5.63	4.64
44	-46	-63	31	5.62	4.48
45	-31	-59	42	5.56	4.45
46	40	-27	18	5.40	5.13
47	55	-7	20	5.28	4.99
48	-12	-15	7	5.10	4.13
49	-4	-33	-3	4.81	4.54
50	55	6	18	4.81	4.40
51	7	-42	48	4.63	2.90
52	-26	-12	53	4.27	3.97
53	56	-17	11	4.24	3.99
54	48	-61	33	4.10	3.85
55	31	-61	-29	4.07	3.82
56	-11	-59	15	3.93	1.61
57	-34	-54	-36	3.84	2.22
58	55	-43	20	3.82	2.88
59	-42	-63	-8	3.81	2.17
60	-51	-38	12	3.41	3.21
61	25	42	30	3.29	1.69
62	-6	44	4	3.27	2.30
63	-24	-33	61	3.24	3.05
64	16	-30	60	3.24	3.00
65	-50	23	19	3.16	2.92
66	41	-3	15	3.15	2.97
67	-52	-24	9	3.13	2.94
68	13	-90	0	3.08	2.22
69	-41	42	6	3.00	2.80
70	-17	23	55	2.62	2.43
71	49	-33	7	2.58	2.43
72	-8	-42	3	2.53	2.35
73	-52	-24	21	2.47	1.61
74	44	33	-9	2.46	1.62
75	40	-48	21	2.45	2.33
76	35	-61	-8	2.45	2.32
77	8	33	24	2.44	2.29
78	9	-56	16	2.40	2.22
79	-58	-42	-3	2.39	1.55
80	-32	-66	-29	2.25	0.74
81	18	-80	-33	1.99	1.88
82	-26	-45	-24	1.75	1.65
83	-39	-38	15	1.75	1.65
84	-53	-24	36	1.74	1.63
85	-51	-51	36	1.72	1.61
86	-46	-21	48	1.71	1.57
87	41	-54	45	1.68	1.58
88	-40	-6	12	1.64	0.86
89	27	-39	-13	1.63	1.54
90	-40	-33	45	1.63	1.54
91	37	-72	12	1.61	1.51
92	8	15	36	1.61	1.49
93	41	-45	9	1.59	1.49
94	-32	-60	-6	1.58	1.45
95	-19	-78	-33	1.56	0.79
96	14	-75	-21	1.55	1.43
97	21	27	49	1.55	1.45
98	-9	-72	37	1.54	1.47
99	40	42	3	1.54	0.73
100	-47	-48	45	1.53	1.41
101	37	12	42	1.52	0.73
102	8	-45	24	1.52	1.39
103	43	22	34	1.51	1.42
104	-2	23	42	1.51	0.73
105	37	36	21	1.49	0.72
106	-28	-87	5	1.47	0.69
107	-3	25	30	0.91	0.86
108	-16	-76	29	0.90	0.84
109	28	-81	12	0.88	0.83
110	42	-73	27	0.87	0.82
111	19	-78	-3	0.87	0.82
112	-40	-42	39	0.80	0.77

Rank	Coordinates X	Y	Z	Age - Procedure 10 FDR	Concensus
39	-6	44	4	32.36	5.94
40	48	18	12	32.30	0.00
41	-53	-24	36	32.16	0.00
42	8	45	21	31.24	3.22
43	-31	-59	42	31.13	0.00
44	40	-27	18	30.20	3.60
45	55	6	18	29.70	5.05
46	41	3	36	28.94	5.64
47	43	-24	45	28.29	0.00
48	-46	-15	35	28.18	0.00
49	-1	10	46	27.88	1.60
50	55	-43	20	27.75	3.45
51	-41	42	6	26.99	1.34
52	-51	-51	36	26.96	0.00
53	-12	-15	7	26.67	7.06
54	-29	-30	9	26.60	0.00
55	31	-15	18	25.95	0.00
56	37	36	21	25.67	2.16
57	-17	-51	0	25.33	2.47
58	-58	-42	-3	24.74	0.00
59	-42	-63	-8	24.73	3.15
60	-8	-42	3	24.49	2.91
61	-46	-63	31	24.47	3.74
62	49	-33	7	24.38	3.37
63	16	-30	60	24.25	0.00
64	8	15	36	24.19	0.00
65	-52	-24	9	24.11	0.00
66	-46	-21	48	24.08	10.58
67	-17	23	55	23.98	6.06
68	-40	-6	12	23.96	0.00
69	-34	-54	-36	23.62	4.06
70	-23	-60	-33	23.30	0.75
71	25	42	30	23.01	11.79
72	8	-45	24	22.86	2.40
73	41	-54	45	22.69	0.00
74	21	27	49	22.40	0.00
75	-35	-69	36	22.19	3.80
76	-11	39	21	22.18	1.99
77	-40	-33	45	22.08	0.00
78	35	-61	-8	21.87	8.56
79	43	-12	26	21.85	0.00
80	-11	-59	15	21.43	3.36
81	-47	-48	45	21.04	0.00
82	56	-17	11	20.93	1.21
83	-3	25	30	20.75	3.25
84	40	42	3	20.44	3.40
85	55	-7	20	20.35	3.66
86	-32	-60	-6	19.61	6.06
87	8	33	24	19.52	0.00
88	37	12	42	19.44	0.00
89	-41	3	36	19.14	7.62
90	32	-73	-29	19.11	3.97
91	-5	-78	-33	18.72	20.14
92	44	33	-9	18.03	0.00
93	-52	-24	21	17.83	0.00
94	31	-61	-29	17.78	0.00
95	27	-39	-13	17.61	10.71
96	-44	6	15	16.69	1.46
97	50	-18	-9	16.62	2.56
98	18	-80	-33	16.59	0.00
99	27	50	23	16.29	23.09
100	37	30	33	16.06	4.65
101	-40	-42	39	15.79	2.00
102	-51	-38	12	15.77	0.39
103	-34	-48	45	15.75	0.09
104	41	-15	39	15.71	10.07
105	13	-90	0	15.64	3.73
106	-39	-38	15	15.52	0.00
107	41	-45	9	15.47	3.45
108	-26	-45	-24	15.44	0.00
109	27	-74	27	15.42	6.06
110	-32	-66	-29	15.40	0.38
111	-16	-76	29	15.16	0.00
112	-5	-45	24	15.12	4.55

(Continued)

Table A1 | Continued

Rank	Coordinates X	Y	Z	Age - Procedure 5 FDR	Concensus	Rank	Coordinates X	Y	Z	Age - Procedure 7 FDR	Concensus	Rank	Coordinates X	Y	Z	Age - Procedure 10 FDR	Concensus
113	32	-73	-29	18.86	2.91	113	-22	-54	-21	0.80	0.73	113	50	-7	34	15.03	0.00
114	31	-61	-29	18.09	1.57	114	-44	6	15	0.80	0.74	114	9	-56	16	14.56	0.00
115	-51	-38	12	17.93	2.20	115	50	-7	34	0.79	0.75	115	-28	-87	5	13.81	4.61
116	-40	-42	39	17.72	1.39	116	-2	-75	29	0.77	0.71	116	-43	-9	48	13.77	0.00
117	-15	-64	-21	17.59	0.00	117	32	27	12	0.77	0.71	117	43	22	34	13.74	0.00
118	-44	6	15	17.43	0.77	118	32	-81	-3	0.76	0.72	118	-26	-43	-11	13.52	10.99
119	-39	-38	15	17.02	2.31	119	-56	-48	9	0.76	0.71	119	7	37	0	13.42	0.00
120	-5	-45	24	16.97	2.14	120	-5	-78	-33	0.76	0.72	120	40	-48	21	12.57	1.32
121	-50	-63	12	16.92	0.69	121	-26	-43	-11	0.75	0.69	121	-43	22	34	12.42	10.51
122	-26	-45	-24	16.29	1.60	122	37	30	33	0.74	0.66	122	-56	-26	-14	12.16	0.00
123	27	50	23	16.04	1.37	123	-43	22	34	0.74	0.65	123	28	-81	12	11.79	0.00
124	-28	-87	5	15.85	0.69	124	-56	-26	-14	0.74	0.69	124	-9	-72	37	11.52	0.00
125	28	-81	12	15.69	0.76	125	-41	3	36	0.74	0.00	125	-50	-63	12	11.21	0.00
126	-53	3	24	15.15	0.70	126	8	-76	12	0.74	0.00	126	-28	-75	24	10.67	0.00
127	-10	-72	-15	14.20	0.69	127	-50	-63	12	0.74	0.00	127	44	-63	5	10.51	0.00
128	-3	25	30	13.45	0.81	128	-53	3	24	0.73	0.00	128	37	-72	12	10.32	10.77
129	-56	-26	-14	12.80	0.00	129	-34	-48	45	0.73	0.00	129	-53	3	24	10.06	5.84
130	-2	23	42	12.63	0.00	130	7	37	0	0.73	0.00	130	-28	51	15	9.76	0.00
131	14	-75	-21	12.29	0.66	131	-5	-45	24	0.72	0.00	131	-22	-54	-21	9.70	3.23
132	-25	45	30	12.22	0.00	132	-5	-80	7	0.00	0.00	132	32	-81	-3	9.49	7.20
133	19	-78	-3	12.05	0.77	133	44	-63	5	0.00	0.00	133	-35	-82	-5	9.20	4.53
134	32	27	12	12.04	1.45	134	-35	-82	-5	0.00	0.00	134	42	-73	27	8.72	1.46
135	-19	-78	-33	11.88	0.76	135	31	-61	-29	0.00	0.00	135	-19	-78	-33	8.53	0.00
136	43	22	34	11.86	0.68	136	-15	-64	-21	0.00	0.00	136	0	-29	30	8.30	0.00
137	-56	-48	9	11.84	0.00	137	27	-74	27	0.00	0.00	137	2	-66	-24	8.00	6.00
138	0	-29	30	11.83	0.00	138	5	-75	-12	0.00	0.00	138	-15	-64	-21	7.53	0.00
139	-28	51	15	11.80	0.00	139	-40	-48	27	0.00	0.00	139	-2	23	42	7.48	1.25
140	-4	-40	43	11.67	0.00	140	26	-90	0	0.00	0.00	140	30	-61	39	7.36	0.00
141	2	-66	-24	11.63	0.00	141	50	-18	-9	0.00	0.00	141	-10	-72	-15	7.27	2.20
142	32	-81	-3	11.58	0.00	142	-23	-60	-33	0.00	0.00	142	-25	45	30	7.16	4.63
143	-22	-54	-21	11.25	0.00	143	-10	-72	-15	0.00	0.00	143	-40	-76	22	7.14	0.00
144	-41	3	36	11.22	0.00	144	2	-66	-24	0.00	0.00	144	14	-75	-21	6.55	3.51
145	42	-73	27	10.87	0.00	145	-4	-93	9	0.00	0.00	145	56	3	36	6.26	0.00
146	-40	-76	22	10.84	0.00	146	-25	45	30	0.00	0.00	146	-32	-66	-29	6.13	0.00
147	30	-61	39	10.74	0.00	147	-11	39	21	0.00	0.00	147	-4	-93	9	6.02	0.00
148	-35	-82	-5	9.67	0.00	148	-5	-60	-15	0.00	0.00	148	52	-33	-15	5.89	0.00
149	-32	-66	-29	8.97	0.68	149	41	-15	39	0.00	0.00	149	32	27	12	5.86	0.00
150	-2	-75	29	8.74	0.72	150	56	3	36	0.00	0.00	150	19	-78	-3	5.79	3.70
151	8	-76	12	7.68	0.00	151	-32	-66	-29	0.00	0.00	151	-2	-75	29	5.37	3.60
152	56	3	36	7.55	0.00	152	-43	-9	48	0.00	0.00	152	8	-76	12	5.17	1.64
153	52	-33	-15	5.64	0.00	153	-28	-75	24	0.00	0.00	153	26	-90	0	4.98	6.33
154	-5	-80	7	4.91	0.00	154	-28	51	15	0.00	0.00	154	-5	-80	7	4.86	0.00
155	-5	-60	-15	4.80	0.00	155	27	50	23	0.00	0.00	155	-4	-40	43	3.76	0.00
156	26	-90	0	4.75	0.00	156	0	-29	30	0.00	0.00	156	-5	-60	-15	3.64	0.00
157	31	-61	-29	3.60	0.00	157	30	-61	39	0.00	0.00	157	-56	-48	9	2.85	0.00
158	5	-75	-12	2.42	0.00	158	-40	-76	22	0.00	0.00	158	31	-61	-29	2.44	0.00
159	-4	-93	9	2.34	0.00	159	-4	-40	43	0.00	0.00	159	5	-75	-12	0.00	11.59
160	-40	-48	27	0.00	0.00	160	52	-33	-15	0.00	0.00	160	-40	-48	27	0.00	0.00

was not exclusionary except for Conduct Disorder, Bipolar and Major Depressive Disorders, as well as any psychotic disorders. Psychostimulant drugs were withheld at least 24 h before scanning. Inclusion criteria for TDC required absence of any Axis-I psychiatric diagnoses per parent and child KSADS-PL interview, as well as T-scores below 60 for all the CPRS-R: LV ADHD summary scales. Estimates of FSIQ above 80, right-handedness and absence of other chronic medical conditions were required for all children.

Peking University/Beijing Normal University

Study participants with the diagnosis of ADHD were initially identified using the Computerized Diagnostic Interview Schedule IV (C-DIS-IV) (Shaffer et al., 2000). Upon referral for participation to the study participation, all participants (ADHD and TDC) were evaluated with the Schedule of Affective Disorders and Schizophrenia for Children—Present and Lifetime Version (KSADS-PL) with one parent for the establishment of the diagnosis for study inclusion. Thus, identification of the ADHD subtype was based on this psychiatric interview. Additional inclusion included: (1) right-handedness, (2) no history of neurological disease and no diagnosis of either schizophrenia, affective disorder, or pervasive development disorder, and (3) full scale Wechsler Intelligence Scale for Chinese Children-Revised (WISCC-R) score of greater than 80. Psychostimulant medications were withheld at least 48 h prior to scanning. Informed consent was also obtained from the parent of each subject and all of the children agreed to participate in the study.

Brown University and Bradley Hospital

Psychiatric diagnoses were based on evaluation by the same board-certified child/adolescent psychiatrist (DPD) for all participants, using the Child Schedule for Affective Disorders Present and Lifetime version (KSADS-PL) administered to parents and children separately (Puig-Antich and Ryan, 1986). All participants completed the WASI as an overall measure of cognitive

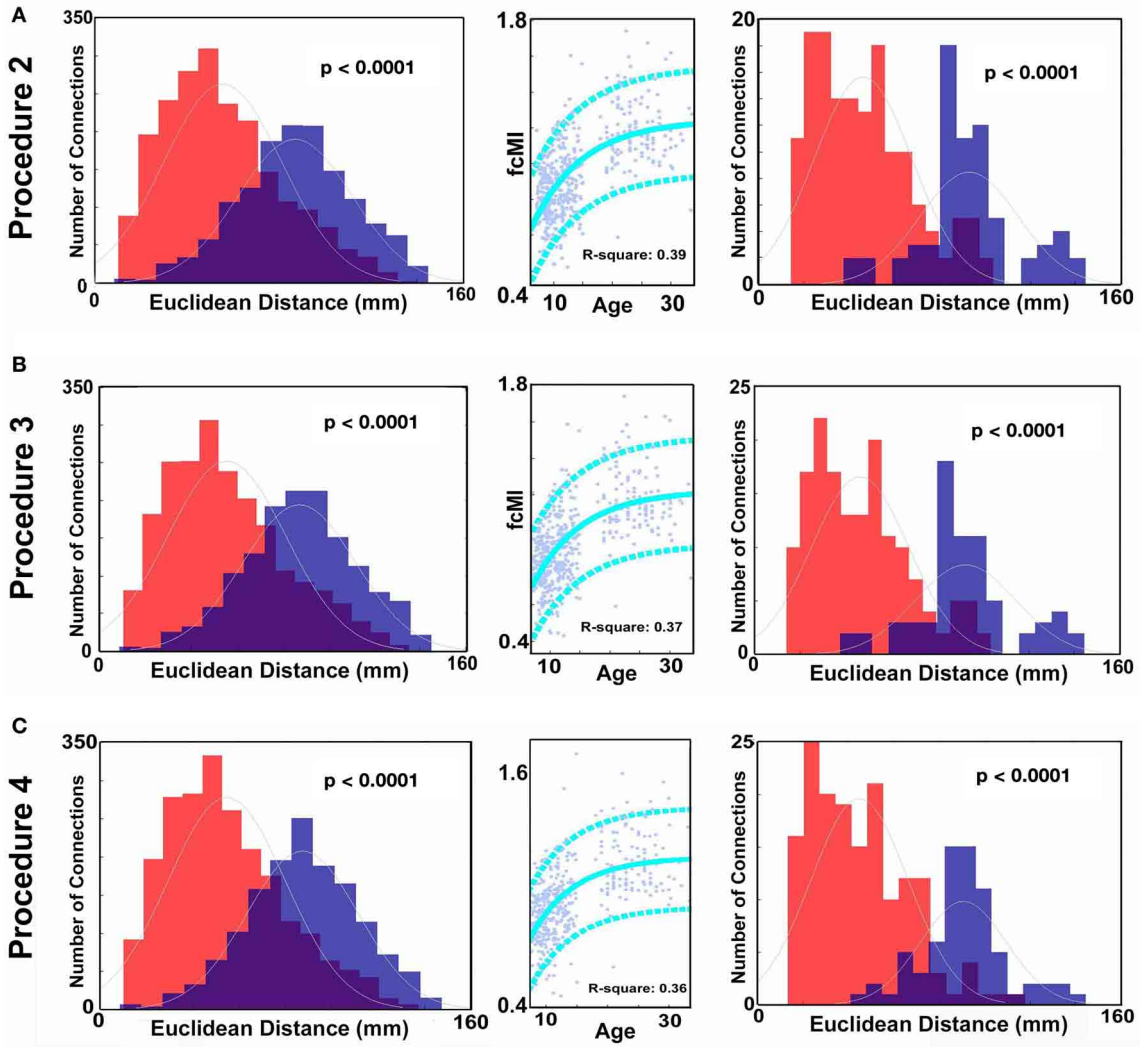

FIGURE A1 | For procedures 2 (A), 3 (B), and 4 (C) (as in Figure 2) on left histograms of Euclidean distance for functional connections that get stronger with age and those that get weaker with age (FDR corrected). As in **Figure 3** (middle) SVR-based MVPA brain maturation curves and on right Euclidean distance of consensus features that grow up with age vs. those that grow down with age.

ability (Wechsler, 1999, 2003). Children in the ADHD group had to meet Diagnostic and Statistical Manual 4th Edition Text Revision (DSM-IV-TR) criteria for ADHD, as determined by parent and child answers to the KSADS-PL and were required to have ongoing psychiatric treatment. Exclusion criteria were comorbid mood or anxiety disorders, any autism spectrum disorder, medical illness that was unstable, or could cause psychiatric symptoms, or substance abuse within ≤2 months of participation. All ADHD participants taking psychostimulant medications (i.e., derivatives of methylphenidate or dextroamphetamine) were scanned when medication-free for five drug half-lives. TDC participant inclusion criteria were a negative history of psychiatric illness in the participant and their first-degree relatives. Exclusion criteria were pregnancy, ongoing medical or neurological illness or past/present psychiatric, or substance disorder. All participants had an IQ greater than 70.

DATA ACQUISITION AND PROCESSING

Oregon Health and Science University

Participants were scanned using a 3.0 Tesla Siemens Magnetom Tim Trio scanner (Siemens, Erlangen, Germany) with a 12-channel head-coil at the OHSU Advanced Imaging Research Center. One high-resolution T1-weighted MPRAGE sequence lasting 9 min and 14 s (TR = 2300 ms, TE = 3.58 ms, orientation = sagittal, 256 × 256 matrix, resolution = 1^3 mm) was collected. Blood-oxygen level dependent (BOLD)-weighted functional imaging data were collected in an oblique plane (parallel to the ACPC) using T2*-weighted echo-planar imaging (EPI) (TR = 2500 ms, TE = 30 ms, flip angle = 90°, FOV = 240 mm, 36 slices covering the whole brain, slice thickness = 3.8 mm, in-plane resolution = 3.8 × 3.8 mm). Steady state magnetization was assumed after five frames (~10 s). Three runs of 3.5 min each were obtained. During rest periods subjects were instructed to

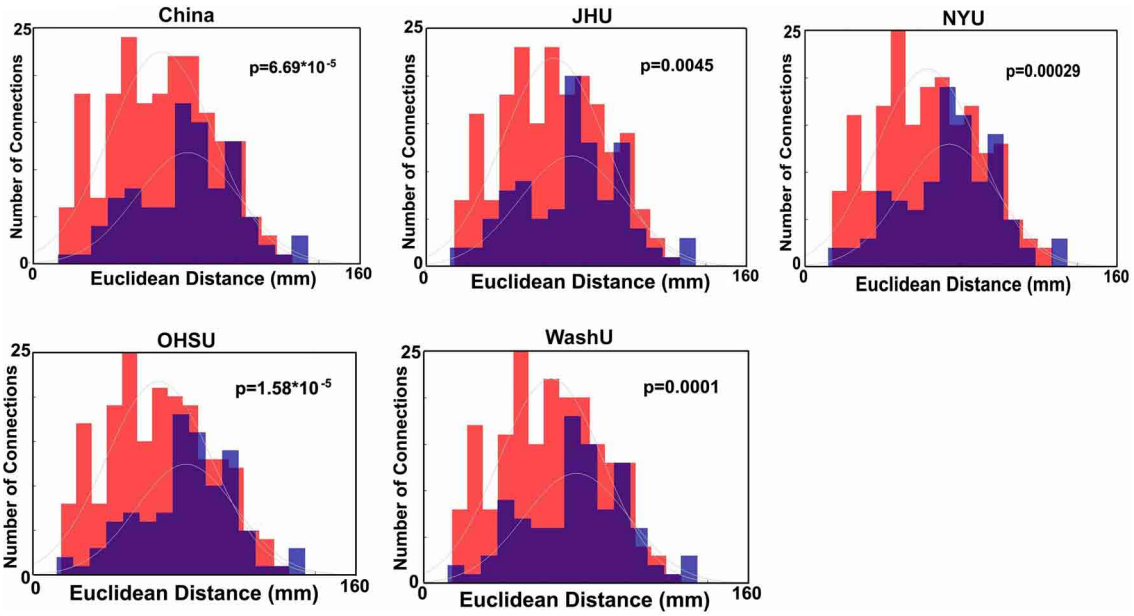

FIGURE A2 | Site-by-site histograms of Euclidean distance for functional connections identified in the whole group analysis that get stronger with age and those that get weaker with age (FDR corrected) using procedure 8.

stay still, and fixate on a standard fixation-cross in the center of the display.

Kennedy Krieger Institute

Participants were scanned using a 3.0 Tesla Philips scanner with an eight-channel head-coil. One high-resolution T1-weighted MPRAGE sequence (TR = 7.99 ms, TE = 3.76 ms, flip angle = 8°) was collected. BOLD-weighted functional imaging data were collected using T2*-weighted EPI (TR = 2500 ms, TE = 30 ms, flip angle = 75°, 2D-SENSE EPI). The run lasted either 5 min 20 s or 6 min 30 s. During rest participants were instructed to relax, stay as still as possible, keep eyes open, and fixate on a center cross.

New York University

Participants were scanned using a Siemens Allegra 3.0 Tesla scanner at the NYU Center for Brain Imaging. For each participant a T1-weighted MPRAGE sequence was acquired using a magnetization prepared gradient echo sequence (TR = 2530 ms; TE = 3.25 ms; TI = 1100 ms; flip angle = 7°; 128 slices; FOV = 256 mm; acquisition voxel size = 1.31 × 1.3 mm). A 6-minute resting scan comprising 180 contiguous whole-brain functional volumes was also acquired for each participant using a multi-echo EPI sequence (TR = 2000 ms; flip angle = 90°; 33 slices; voxel size = 3 × 3 × 4 mm; effective TE = 30 ms, FOV = 240 × 192 mm). During rest periods participants were instructed to lie still and relax with their eyes open, while a standard fixation-cross was presented in the center of the display.

Peking University samples

Dataset #1. Images were acquired using a Siemens Trio 3.0 Tesla scanner in National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University. For each participant,

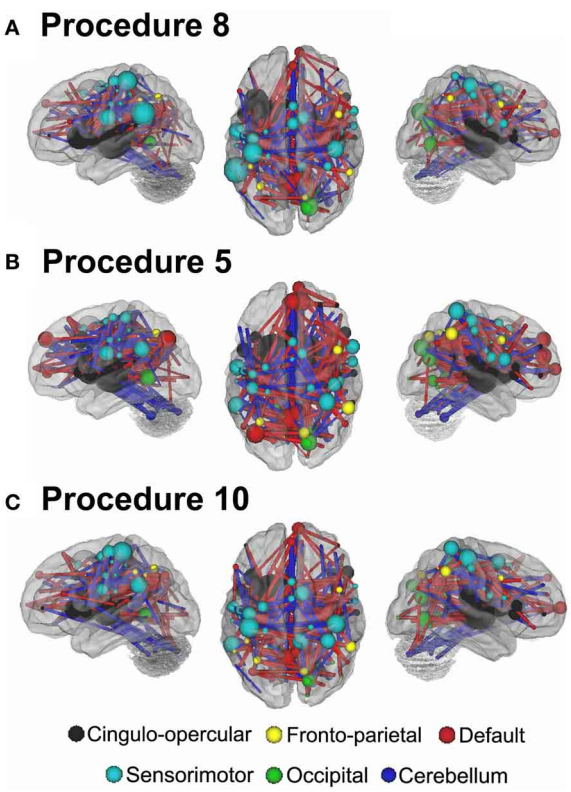

FIGURE A3 | As with Figure 7, Node strength for the consensus features that best predict age for Procedure 8 in (A), Procedure 5 in (B), and Procedure 10 (C). Included are the vectors, which are used to calculate node strength.

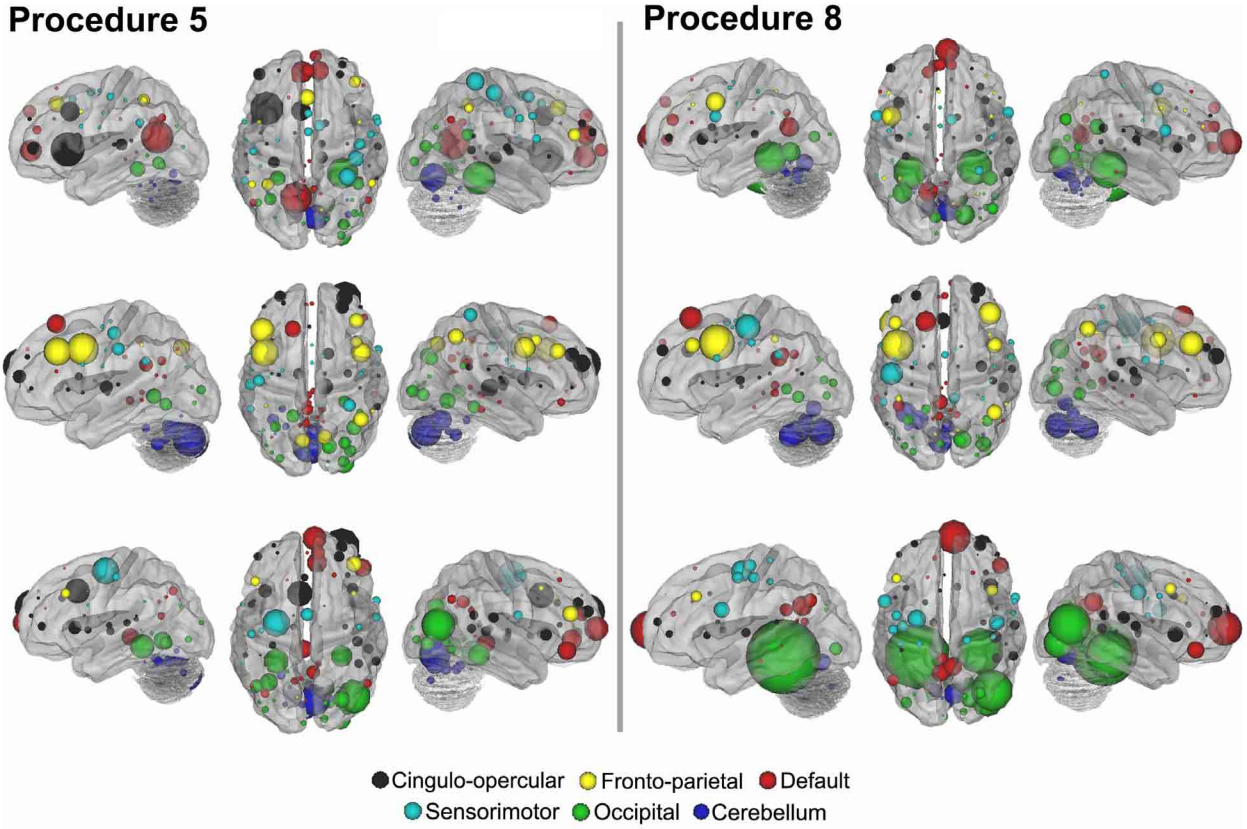

FIGURE A4 | *Post-hoc* comparisons on the consensus features from the 2-group classification for Procedures 5 and 8. While not identical, patterns across all of these movement correction procedures is largely consistent. **Top:** TDC vs. ADHD-C. **Middle:** TDC vs. ADHD-I. **Bottom:** ADHD-C vs. ADHD-I. **Left:** Procedure 5. **Right:** Procedure 8 [Node colors represent network categorization stemming from a community detection procedure performed for a previous report (Dosenbach et al., 2010). Red—default; blue—cerebellum; yellow—fronto-parietal; black—cingulo-opercular; green—occipital; cyan—sensorimotor].

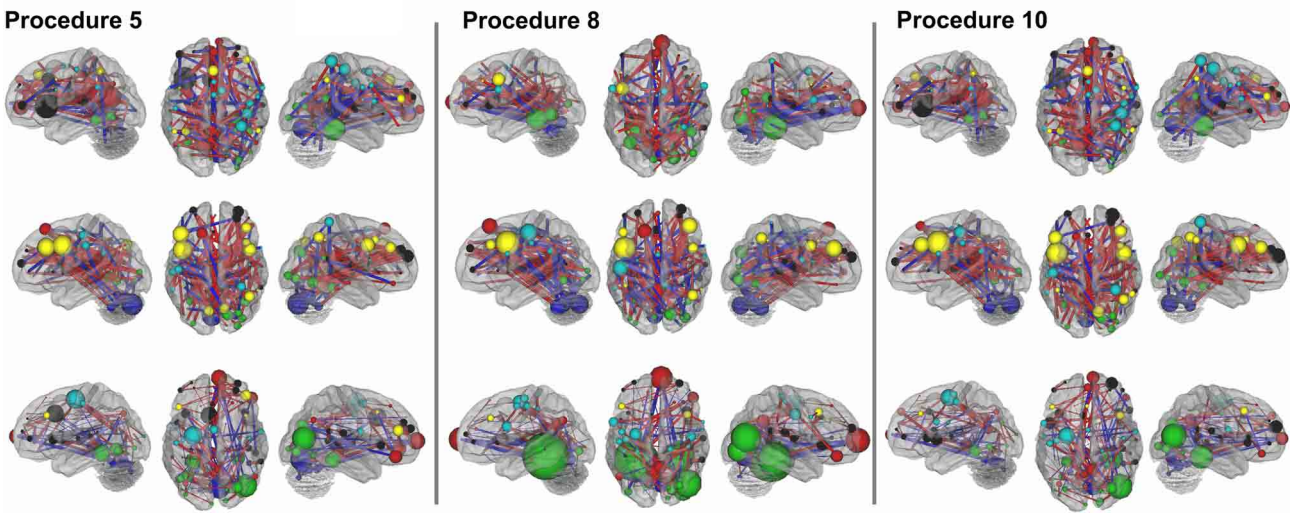

FIGURE A5 | As in Figures 8 and A4, *Post-hoc* comparisons on the consensus features from the 3-group classification. Included are the vectors, which are used to calculate node strength. **Top:** TDC vs. ADHD-C. **Middle:** TDC vs. ADHD-I. **Bottom:** ADHD-C vs. ADHD-I. **Left:** Procedure 5. **Middle:** Procedure 8. **Right:** Procedure 10.

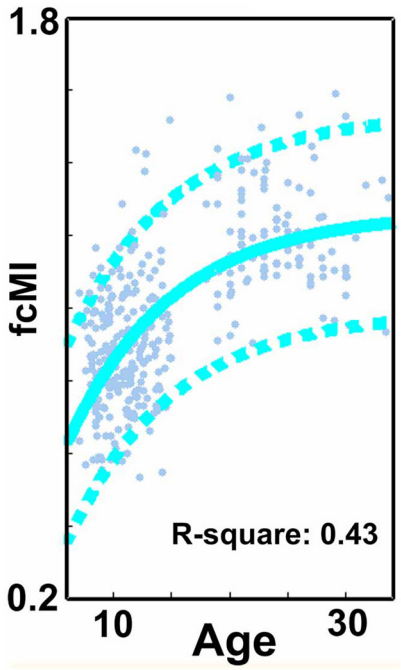

FIGURE A6 | A re-examination of SVR-based MVPA brain maturation curve for Procedure 7 (frame removal) while including only those subjects with <60% of their frames removed.

a high-resolution T1-weighted anatomical image was acquired [128 sagittal slices, slice thickness/gap = 1.33/0 mm, in-plane resolution = 256 × 192, TR = 2530 ms, TE = 3.39 ms, inversion time (TI) = 1100 ms, flip angle = 7°, FOV = 256 × 256 mm^2]. A resting-state scan was obtained for each participant (33 axial slices, TR = 2000 ms, TE = 30 ms, flip angle = 90°, thickness/gap = 3.5/0.7 mm, FOV = 200 × 200 mm^2, matrix = 64 × 64, 240 volumes), as well as diffusion tensor imaging (not reported here).

Dataset #2. Images were acquired using a Siemens Trio 3.0 Tesla scanner in National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University. All of the resting-state functional data were acquired using an EPI sequence with the following parameters: 33 axial slices, TR = 2000 ms, TE = 30 ms, flip angle = 90°, slice thickness/skip = 3.0/0.6 mm, FOV = 200 × 20 mm, in-plane resolution = 64 × 64, 240 volumes. For each patient, T1-weighted structural images were acquired using a spoiled gradient-recalled sequence covering the whole brain and used for the purpose of image registration (see Data preprocessing). The T1-weighted structural images were acquired with the following parameters: 176 sagittal slices, TR = 2530 ms, TE = 3.45 ms, flip angle = 7°, slice thickness/skip = 1.0/0 mm, FOV = 256 × 256 mm, in-plane resolution = 256 × 256.

Children with ADHD were scanned twice, in a double-blinded, randomized, counterbalanced way. The two scans were at least 2 days apart, and each scan was taken 1 h after either 10 mg MPH administration or placebo (Vitamin B6, 10 mg). All the patients

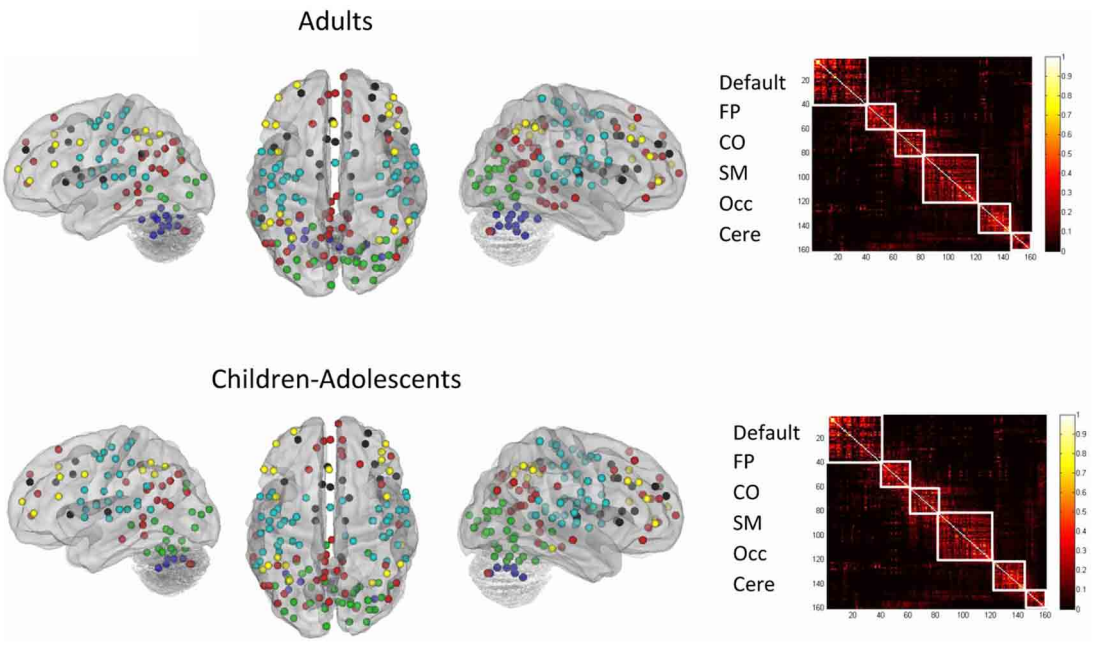

FIGURE A7 | Network categorization of each ROI was based on labels designated in a previous report (Dosenbach et al., 2010). This categorization stemmed from a community detection procedure conducted on combined correlation matrices across adult subjects [e.g., see Fair et al. (2009)]. The modularity optimization algorithm of Newman was used (Newman, 2006). The modules (i.e., communities) used to categorize regions had a high quality index (Q) and were the most resistant to perturbation by randomization, measured by variation of information (VOI) (Karrer et al., 2008). These community assignments were re-examined here for both children and adults after applying procedure 8. In this instance, the weight conserving community detection algorithm used was based on the work by Rubinov and Sporns (2011). Network assignments largely agreed with the previous report and were similar in children and adults. On the left network assignments are visualized on the brain and on the right in matrix form.

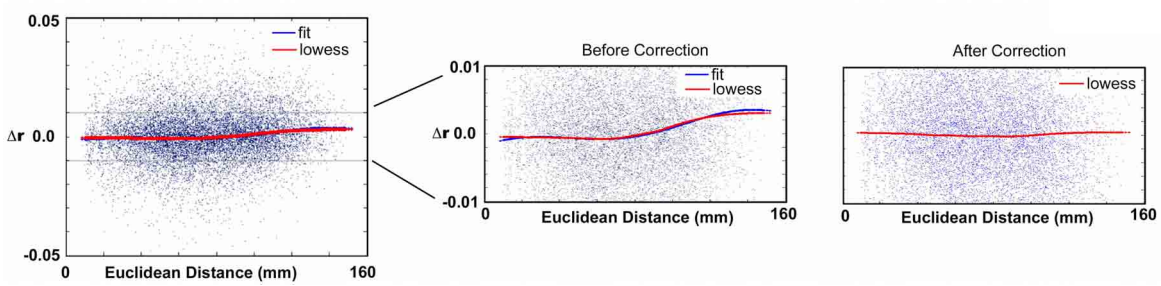

FIGURE A8 | Sample corrected subject data for Procedure 10. For a given subject, correlation matrices are re-created after censoring (i.e., scrubbing). Differences in the correlation values from the original matrix and the censored matrix are then plotted as a function of distance. The red curve shows the trend of delta *r* based on LOWESS smoothing. The blue curve shows the polynomial (see Methods) fit. The method is applied on each subject individually such that every subject will have a fit for their respective curve. The fit is then used to regress the effect of movement on the original *r*-values. Censoring the frames as in Procedure 7 after applying this procedure reveals no effect of movement (Right panel).

had not received stimulant treatment for at least 2 days before the first scan, and were asked not to take any stimulant between two scans. The control boys were scanned once without MPH or placebo taken for ethical reasons. Only placebo scans were used for the present study.

Dataset #3. This dataset was previously employed in a prior study (Cao et al., 2009).

Images were acquired using a Siemens Trio 3.0 Tesla scanner in the Institute of Biophysics, Chinese Academy of Sciences. Rest scans were acquired using an EPI sequence with the following parameters: 30 axial slices, TR = 2000 ms, TE = 30 ms, flip angle = 90°, thickness/skip = 4.5/0 mm, FOV = 220 × 220 mm, matrix = 64 × 64, 240 volumes. Participants were asked simply to remain still, close their eyes, think of nothing systematically, and not fall asleep. Additionally, for each participant, a high-resolution T1-weighted anatomical image using a spoiled gradient-recalled sequence covering the whole brain was acquired. The data were collected in a period of about 2 years and some modifications were made in the sequence of the structural images. Most of the subjects (see details below) were scanned with one of the following two kinds of parameters: (1) 192 slices, TR = 2000 ms, TE = 3.67 ms, inversion time = 1100 ms, flip angle = 12°, FOV = 240 × 240 mm, matrix = 256 × 256, used in 8 patients and 12 controls; (2) 176 slices, TR = 1770 ms, TE = 3.92 ms, inversion time = 1100 ms, flip angle = 12°, FOV = 256 × 256 mm, matrix = 512 × 512, used in 9 patients and 11 controls. Other scanning sessions, which have no relation to the present study, are not described here.

Brown University

Scans were acquired on a Siemens Tim Trio 3.0 Tesla scanner with a 12-channel head-coil. A high-resolution T1-weighted MPRAGE anatomical image was acquired for normalization and localization (TR = 2250 ms, TE = 2.98 ms, T1 = 900 ms, flip angle = 9°, slices = 160, FOV = 256 mm, voxels = 1 × 1 × 1 mm). The resting-state functional connectivity scan contained 256 continuous BOLD volumes (TR = 2000 ms, TE = 25 ms, flip angle = 90°, slices = 35, FOV = 192 mm, voxels = 3 × 3 × 3 mm). The scan lasted for 8 min and 36 s. During the scan, participants were instructed to rest with their eyes open while the word "relax" was back-projected via LCD projector.

Washington University

Scans were acquired on a Siemens Tim Trio 3.0 T Scanner with a Siemens 12-channel Head Matrix Coil. A T1-weighted sagittal MP-RAGE anatomical image was acquired (TE = 3.06 ms, TR-partition = 2.4 s, TI = 1000 ms, flip angle = 8°, 176 slices with 1 × 1 × 1 mm voxels). The resting-state functional connectivity scan were obtained using a BOLD contrast sensitive gradient echo echo-planar sequence (TE = 27 ms, flip angle = 90°, in-plane resolution = 4 × 4 mm; volume TR = 2.5 s). Whole-brain coverage for the functional data was obtained using 32 contiguous interleaved 4 mm axial slices.

REFERENCES

Cao, X., Cao, Q., Long, X., Sun, L., Sui, M., Zhu, C., et al. (2009). Abnormal resting-state functional connectivity patterns of the putamen in medication-naive children with attention deficit hyperactivity disorder. *Brain Res.* 1303, 195–206.

Conners, C. K. (2008). *Conners 3rd Edition Manual.* Toronto, ON: Multi-Health Systems Inc.

DuPaul, G., Power, T., Anastopoulos, A., and Reid, R. (1998). *ADHD Rating Scales-IV: Checklists, Norms and Clinical Interpretation.* New York, NY: Guilford Press.

Epstein, J. N., Conners, C. K., Erhardt, D., March, J. S., and Swanson, J. M. (1997). Asymmetrical hemispheric control of visual-spatial attention in adults with attention deficit hyperactivity disorder. *Neuropsychology* 11, 467–473.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J. Am. Acad. Child Adolesc. Psychiatry* 36, 980–988.

Puig-Antich, J., and Ryan, N. (1986). *The Schedule for Affective Disorders and Schizophrenia for School-Age Children*. Pittsburg, CA: Western Psychiatric Institute and Clinic.

Reich, W., Welner, Z., and Herjanic, B. (1997). *Diagnostic Interview for Children and Adolescents-IV (DICA-IV)*. North Tonawanda, NY: Multi-Health Systems, Inc.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., and Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version, IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *J. Am. Acad. Child Adolesc. Psychiatry* 39, 28–38.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence, 3rd Edn*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children Technical and Interpretive Manual, 4th Edn*. San Antonio, TX: The Psychological Corporation.