Supplementary Online Content

Supplement 1

Karalunas SL, Fair D, Musser ED, Aykes K, Iyer SP, Nigg JT. Subtyping attention-deficit/hyperactivity disorder using temperament dimensions: toward biologically based nosologic criteria. *JAMA Psychiatry*. Published online July 9, 2014. Retracted and replaced March 7, 2018. doi:10.1001/jamapsychiatry.2014.763

Supplement. Replacement supplement with errors highlighted (eAppendix, eReferences, eTable 1, eTable 2, eTable 3, eFigure)

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix. Methods and results

I. Diagnostic Procedures

After an initial screening phone call, children were identified for the study via a multi-gate, best-estimate confirmation procedure. In the multi-gate process, a parent/guardian and teacher completed standardized rating scales, including Child Behavior Checklist/Teacher Report Form Attention Problems subscale [CBCL/TRF, 1], Conners' Rating Scales-Revised [CRS-R, 2], and the ADHD Rating Scale [ADHD-RS, 3]. The parent/guardian also completed a semi-structured clinical interview administered by a Master's-level clinician who had achieved research reliability on the interview [Kiddie Schedule for Affective Disorders and Schizophrenia, K-SADS, 4]. Parents rated their child's overall level of impairment in academic, family, and peer relationships using the Strengths and Difficulties Questionnaire (SDQ). IQ was estimated based on a reliable and valid three-subtest short form of the WISC-IV [Vocabulary, Block Design, and Information, 5].

Then, in the best estimate procedure, final diagnoses were made by a clinical diagnostic team (a board certified child psychiatrist and licensed clinical child psychologist), who took into account data from the parent and teacher ratings, parent clinical interview, IQ and achievement testing, and behavioral observations. In both cohorts, agreement ratings between the members of the diagnostic team were acceptable for ADHD diagnosis (kappa>.74) and for other disorders with >5% base rate in the sample (all kappas >.68). Disagreements were conferenced. If consensus was not readily achieved the case was excluded. For all analyses, comorbid diagnoses from the diagnostic team were collapsed into the following categories: 1) "Mood Disorder" (including a diagnosis of either Major Depressive Disorder or Dysthymia); 2) "Anxiety Disorder" (including Generalized Anxiety Disorder, Separation Anxiety Disorder, Social Anxiety Disorder, Obsessive-Compulsive Disorder, or Panic Disorder); 3) Oppositional Defiant

Disorder (including conduct disorder); and 4) "Any Disorder" if they met criteria for any of the three other disorder categories.

Children were excluded if they: were prescribed long-acting psychotropic medications; had neurological impairment, seizure history, head injury with loss of consciousness, other major medical conditions, or substance abuse disorder; had prior diagnosis of mental retardation, autism spectrum disorder, or psychosis; were experiencing a major depressive episode at the time of diagnostic interview; or had estimated IQ <70. Children with ADHD taking stimulant medications (37%) were included in the study. Parents rated children as if not taking medication.

Eight children with ADHD-Hyperactive presentation were excluded because of questions about validity of this category [6], which is rare rare in school-age children [7], and to facilitate the identification of subgroups in a sample considered homogeneous from a diagnostic perspective. The final N=247 children with ADHD is after excluding these 8 children.

II. Temperament Measures

The TMCQ items combine in to 16 scales based on prior factor analysis [8]: Activity Level, Affiliation, Anger/Frustration, Assertiveness/Dominance, Attention Focusing, Discomfort, Fantasy/Openness, Fear, High Intensity Pleasure, Impulsivity, Inhibitory Control, Low Intensity Pleasure, Perceptual Sensitivity, Sadness, Shyness, and Soothability/Falling Reactivity. (The Activation Control scale was omitted from the current analyses because it is an experimental scale on this measure.) Scale reliabilities in our sample (Cronbach's α) ranged from 0.71 to 0.94. Items on the Activity Level, Attention Focusing, Impulsivity, and Inhibitory Control scales overlap with DSM-5 ADHD presentations and this was taken into account in the interpretation.

III. Definition of Quality Index (Q)

Community detection was applied to the child matrix using Rubinov and Sporns' [9] weight-conserving modularity algorithm. In this case, Q is a weighted combination of the modularity for positive relationships and negative relationships defined formally as:

$$Q^* = Q^+ + \frac{v^-}{v^+ + v^-} Q^-$$
$$= \frac{1}{v^+} \sum_{ij} \left(w_{ij}^+ - e_{ij}^+\right) \delta_{M_i M_j} - \frac{1}{v^+ + v^-} \sum_{ij} \left(w_{ij}^- - e_{ij}^-\right) \delta_{M_i M_j}$$

The asymmetric modularity Q*, is a combination of the modularity measure defined in Newman (2006) for the positive links in a graph (Q^+) with a weighted measure of the modularity of the negative links (Q^-). The connection weight between nodes i and j is represented by w_{ij}, where positive and negative weights can be decomposed as w_{ij}^+ and w_{ij}^- respectively. The chance-expected within module weights, which represent the null model of the system, is the product of the positive or negative strengths for nodes i and j, i.e., $s_i^{\pm}=\Sigma_j w_{ij}^{\pm}$, divided by the total weight, $v^{\pm}=\Sigma_{ij} w_{ij}^{\pm}$, i.e. $e_{ij}^{\pm} = \frac{s_i^{\pm} s_j^{\pm}}{v^{\pm}}$. $\delta_{M_i M_j}$is a binary term equal to unity when nodes i and j are in the same module, and zero when they are in different modules. The total weight, referred to by the terms $v^{\pm}$ in the above equation, are defined as the sum of strengths of all positive or negative connections (counted twice as w_{ij} and w_{ji}). The negative weights are explicitly weighted to reduce their contribution to modularity, such that high values of Q* should theoretically agree with more positive weights within modules and negative weights between modules. Conceptually, *Q* represents the overall segregation between identified communities, with higher values indicating stronger separation of communities.

IV. Physiological Recording

Briefly, children watched four, two-minute film clips while physiological recordings were made. The first two clips elicited negative emotions, while the last two segments elicited

positive emotions. In the induction condition, children were asked to facially mimic the emotion of the main character. In the suppression condition, children were instructed to imagine what the main character was feeling, but to keep his or her face still, masking (suppressing) the emotion. A neutral baseline period of two minutes during which children observed a set of ten neutral pictures from the International Affective Picture System [IAPS, 10] was presented prior to the task, accounting for the physiological response of orienting and attending [11].

Physiological recordings were made using disposable silver/silver-chloride electrodes placed in a modified Lead-II electrocardiogram (ECG) configuration with four additional electrodes for impedance cardiography (ICG). ECG and ICG were recorded throughout each of the baselines and task epochs. The R-R series was sampled at 1000 Hz. Interbeat-interval (IBI) and respiration rate (RR) data were derived using the ECG and ICG data.

For cardiac PEP, artifacts were examined and removed using MindWare Impedance Cardiography V.2.6 software [12], allowing for simultaneous editing of the data obtained from ECG and ICG. Visual inspection of the data was completed by two raters (k >0.85 for each epoch). There were no between-group differences in the rate of artifacts (all p >0.50).

For RSA, R-R waves were examined for artifacts and outliers using MindWare® Heart Rate Variability software V.2.6 [13]. Artifacts were removed using the software and visual inspection completed by two raters for validity (all k >0.91). There were no between-group differences in the rate of artifacts (all p >0.50). RSA was indexed by extracting the high frequency component (>0.15 Hz) of the R-R peak time series in 60 second epochs. The time series was detrended and submitted to a Fourier transformation. The high frequency band ($\ln[ms^2]$) was set over the respiratory frequency band of 0.24 to 1.040 Hz. Respiratory rates and amplitudes were derived from the impedance cardiograph signal (Z0) ensuring that these signals

remained within the analytical bandwidth.

V. fMRI

One high-resolution T1-weighted MPRAGE sequence (orientation=sagittal, TE=3.58 ms, TR=2300ms, 256x256 matrix, resolution=1^3mm, total scan time=9 min 14sec) was collected. Blood-oxygen level dependent (BOLD)-weighted functional imaging was collected in an oblique plane (parallel to the ACPC) using T2*-weighted echo-planar imaging (TR=2500ms, TE=30ms, flip angle=90°, FOV=240mm, 36 slices covering the whole brain, slice thickness =3.8mm, in-plane resolution=3.8 x 3.8mm). Steady state magnetization was assumed after 4 frames (~10 s). During rest children were instructed to stay still and fixate on a standard fixation-cross in the center of the display.

Data preprocessing. Data were processed to remove artifacts. These steps included: (i) removal of a central spike caused by MR signal offset, (ii) correction of odd vs. even slice intensity differences attributable to interleaved acquisition without gaps, (iii) correction for head movement within and across runs, and (iv) within-run intensity normalization to a whole brain mode value of 1,000. Atlas transformation of the functional data was computed for each individual via the MPRAGE scan. Each run then was resampled in atlas space [14] on an isotropic 3mm grid, combining movement correction and atlas transformation in one interpolation [15].

Connectivity preprocessing followed prior methods [16, 17]. These steps included: (i) a temporal band-pass filter (0.009 Hz <f <0.08 Hz) and spatial smoothing (6mm full width at half maximum), (ii) regression of six parameters obtained by rigid body head motion correction, (iii) regression of the whole brain signal averaged over the whole brain, (iv) regression of ventricular signal averaged from ventricular region of interest (ROI), and (v) regression of white matter

signal averaged from white matter ROI. Regression of first order derivative terms for the whole brain, ventricular, white matter, and movement signals were also included in the correlation preprocessing.

Motion Correction Procedures. Motion correction was additionally approached in several ways. First, data were screened using an analysis of head position based on rigid body translation and rotation calculated as root mean square (RMS). Total RMS values were calculated on a run-by-run basis for each participant. Participant's BOLD runs with an excess of movement 1.5mm RMS were removed. Second, motion correction using framewise displacement was conducted. Framewise displacement measures movement of any given frame relative to the previous frame. 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 $[d_{ix},d_{iy},d_{iz},\alpha_i,\beta_i,\gamma_i]$. Frames were removed prior to creating the functional connectivity maps for each subject based on an $FD_i>0.3$mm. Two frames after the censored frame, and one frame before were also removed. Subjects with more than 50% of their frames removed were not included in the analyses ($n_{Control}=1$, $n_{ADHD}=16$). This procedure is described in detail in [14].

Sixteen children with ADHD and 1 typically-developing child were screened out due to excessive head motion, leaving a final sample of 39 children with ADHD and 15 controls for analysis. Rates of exclusion based on motion did not differ between the temperament types. Excluded children did not differ from included children in age, FSIQ, symptom count, or parent-reported impairment (all $p>.20$).

VI. Longitudinal Follow-up

Children who were followed did not differ significantly from the children who were not

yet followed in FSIQ, baseline severity (number of symptoms and ratings of impairment), or comorbidity (all $p>.216$), but were younger at baseline than children who were not followed. The age effect was by design because recruitment for the on-going longitudinal study began with younger children and then proceeded to older children.

VII. Results

Full Sample. We first examined the full sample of 437 children and identified three profiles of children ($Q=.55$) with sizes of 192, 181, 64 . As described in the main text, Profile 1 included 82% of the control children and only 13% of the ADHD children. Profiles 2 and 3 together included 87% of the ADHD children and only 18% of non-ADHD controls. We therefore concluded that our initial clustering analysis recovered primarily control and ADHD groups. However, to ensure that this separation of ADHD and typically-developing children was not attributable to differences in age or gender between the samples, we created age-residualized and gender-residualized scores for each of the temperament scales and re-ran the community detection analysis in the full sample two additional times. Using the age- or gender-residualized scores also identified 3 groups with fit indices of Q=.55 and Q=.56, respectively. Results were highly similar to the original analysis. Profile 1 captured 81% (age-residuals) or 82% (gender-residuals) of the non-ADHD controls, while Profiles 2 & 3 captured 86% (age-residual) or 88% (gender-residuals) of the ADHD sample. Together, results reassured us that the original separation of the ADHD and non-ADHD samples was not attributable to age or gender, and we proceeded to the second step of our sequential analysis plan, applying community detection within typically-developing and ADHD groups separately.

Non-ADHD Sample. Community detection using the 190 typically-developing control children identified three profiles with sizes of 21, 98, and 71; however, the quality quotient

indicated only marginal separation of the groups (Q=.21). VOI analyses also indicated that the subgroups in the control population were not as robust as identified in the ADHD sample (eFigure 1). Together, results indicated that temperament subgroups in the non-ADHD sample were marginal to non-existent, and thus we proceeded to focus our analysis on the ADHD sample as described in the main text.

We note in prior work, using neuropsychological data, robust and similar sub-populations were identified in children and with and without ADHD [18]. In contrast, for the current study using temperament ratings, temperament types were only weakly identified in the typically-developing sample. This difference across domains highlights the importance of input feature selection in grouping analyses because any sub-populations identified will be fully determined by the input features. One possibility is that while cognitive subgroups exist in the typically-developing population, temperament subgroups do not; however measurement characteristics of neuropsychological versus temperament measure also may merit more scrutiny to fully explore this picture. Precisely because different groups may be found with different input features, the utility of any subtyping effort is determined via external validation as demonstrated here in the ADHD sample.

eReferences

1. Achenbach, T.M. and L.A. Rescorla, *Manual for the aseba school-age forms & profiles*. 1991, Burlington, VT: Research Center for Children, Youth, & Families.
2. Conners, C.K., *Conners' rating scales: Revised technical manual*. 2003, New York, NY: Multi-Health Systems.
3. DuPaul, G., et al., *ADHD Rating Scale—IV: Checklists, Norms, and Clinical Interpretation*. 1998, NY, NY: Guilford Press.
4. Kaufman, J., et al., *Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime version (K-SADS-PL): Initial reliability and validity*. 1997, Pittsburgh: Department of Psychiatry, University of Pittsburgh School of Medicine.
5. Wechsler, D., *Wechsler Intelligence Scale for Children, 4th Ed (WISC-IV) Technical and Interpretive Manual*. 2003, San Antonio: Harcourt Brace.
6. Willcutt, E.G., et al., *Validity of DSM–IV Attention Deficit/Hyperactivity Disorder Symptom Dimensions and Subtypes.* Journal of Abnormal Psychology, 2012. **121**(4): p. 991-1010.
7. Wolraich, M.L., et al., *Examination of DSM-IV criteria for attention deficit/hyperactivity disorder in a county-wide sample.* Journal of developmental & behavioral pediatrics, 1998. **19**(3): p. 162.
8. Simonds, J. and M.K. Rothbart, *The Temperament in Middle Childhood Questionnaire (TMCQ): A computerized self-report measure of temperament for ages 7- 10*, in *Occasional Temperament Conference*. 2004: Athens, GA.
9. Rubinov, M. and O. Sporns, *Weight-conserving characterization of complex functional brain networks.* Neuroimage, 2011. **56**(4): p. 2068-2079.
10. Lang, P.J., M.M. Bradley, and B.N. Cuthbert, *International affective picture system (IAPS): Technical manual and affective ratings*. 1999, Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.
11. Porges, S.W., *Orienting in a defensive world: Mammalian modifications of our evolutionary heritage. A polyvagal theory.* Psychophysiology, 1995. **32**(4): p. 301-318.
12. *Mind Ware Impedance Cardiography*. 2008, MindWare Technologies: Gahanna, OH.
13. *Mind Ware Heart Rate Variability*. 2008, MindWare Technologies: Gahanna, OH.
14. Talairach J, Tournoux P. (1988) *Co-Planar Stereotaxic Atlas of the Human Brain*. Stuttgart: Thieme.
15. Lancaster, J.L., et al., *Automated Talairach atlas labels for functional brain mapping.* Human Brain Mapping, 2000. **10**(3): p. 120-131.
16. Fair, D.A., et al., *Maturing thalamocortical functional connectivity across development.* Frontiers in Systems Neuroscience, 2010. **4**.
17. Fox, M.D., et al., *The human brain is intrinsically organized into dynamic, anticorrelated functional networks.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(27): p. 9673-9678.
18. Fair, D., et al., *Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **10.1073/pnas.1115365109**.

eTable 1. *Sample characteristics.*

	Control	***ADHD***	***F(2, 439)***
Basic Demographics			
N	190	247	
(Boys:Girls)	88:102	177:70	$\chi^2(2)=27.34^{***}$
Full Scale IQ	114.1 (13.2)	106.4 (14.9)	10.99^{***}
Age (months)	100.0 (13.2)	110.0 (15.7)	11.59^{***}
% on Stimulant Medications	---	36.8%	
Mood Disorder (% of group)	2.6%	10.1%	$\chi^2(2)=9.67^{**}$
Anxiety Disorder (% group)	14.1%	28.3%	$\chi^2(2)=18.80^{***}$
ODD (% of group)	3.1%	27.5%	$\chi^2(2)=46.19^{***}$
Any Disorder (% of group)	17.2%	51.0%	$\chi^2(2)=54.00^{***}$
Temperament Scales			
Activity Level	3.87 (0.67)	4.04 (0.71)	7.09^{**}
Affiliation	4.39 (0.47)	4.18 (0.53)	19.07^{***}
Anger	2.62 (0.63)	3.31 (0.78)	97.41^{***}
Discomfort	2.45 (0.67)	2.73 (0.72)	18.42^{***}
Fear	2.39 (0.69)	2.51 (0.74)	3.23
Openness	4.32 (0.52)	4.11 (0.57)	15.75^{***}
High Intensity Pleasure	3.41 (0.57)	3.72 (0.61)	29.60^{***}
Impulsivity	2.44 (0.56)	3.66 (0.67)	409.42^{***}
Inhibitory Control	3.71 (0.55)	2.61 (0.55)	430.67^{***}
Sadness	2.41 (0.61)	2.74 (0.69)	24.95^{***}
Shyness	2.55 (0.79)	2.53 (0.92)	0.04
Soothability	3.77 (0.71)	3.22 (0.70)	62.74^{***}
Assertiveness/Dominance	3.63 (0.49)	3.76 (0.57)	6.11^{*}
Attentional Focus	3.95 (0.61)	2.15 (0.58)	997.90^{***}
Low Intensity Pleasure	3.97 (0.52)	3.52 (0.59)	68.32^{***}
Perceptual Sensitivity	3.52 (0.62)	3.34 (0.70)	8.55^{**}

eTable 2. *Characteristics of Temperament Types in ADHD Sample.*

	Type 1 **"Mild"**	***Type 2*** **"Surgent"**	***Type 3*** ***"Irritable"***	***F(2, 244)***	***Post-hoc***
Basic Demographics					
(Boys:Girls)	25:39	67:18	71:27	$\chi^2(2)=5.80$	
DSM-V Presentations (I:C)	36:28	8:77	21:77	$\chi^2(2)=43.31^{***}$	
Full Scale IQ	111.6 (14.4)	106.7 (13.2)	107.1 (14.7)	2.58	
Age (months)	109.4 (16.6)	101.7 (15.1)	104.7 (14.8)	4.58^{*}	1 > 2
% on Stimulant Medications	24.3%	34.1%	47.3%	$\chi^2(2)=5.21$	
Inattention Symptoms	7.9 (1.0)	8.3 (1.2)	8.4 (0.8)	3.87^{*}	1 < 3
Hyper-Impulsive Symptoms	5.7 (2.7)	7.6 (1.4)	7.1 (1.9)	15.38^{***}	1 < 2, 3
T1 SDQ Impairment	0.7 (0.9)	1.0 (0.9)	1.5 (0.8)	15.80^{***}	1, 2 < 3
Mood Disorder (% group)	7.8%	10.6%	11.2%	$\chi^2(2)=0.53$	
Anxiety Disorder (% group)	28.1%	21.2%	41.8%	$\chi^2(2)=9.43^{**}$	
ODD (% of group)	12.2%	23.6%	41.8%	$\chi^2(2)=19.26^{***}$	
Any Disorder T1 (% group)	39.1%	43.5%	65.3%	$\chi^2(2)=13.57^{**}$	
Any New Disorder Onset at T2 (% of group)[a]	18.2%	11.1%	39.5%	$\chi^2(2)=9.15^{*}$	
Temperament Scales					
Activity Level	-.37 (1.12)	.85 (.72)	.16 (1.00)	31.23^{***}	2 > 3 > 1
Affiliation	-.42 (1.01)	.07 (.95)	-.93 (1.13)	20.91^{***}	2 > 1 > 3
Anger	.14 (1.07)	.78 (1.03)	2.00 (.87)	77.21^{***}	3 > 2 > 1
Discomfort	.29 (.77)	-.24 (.90)	1.09 (.99)	50.41^{***}	3 > 1 > 2
Fear	.05 (.75)	-.42 (.93)	.78 (1.06)	38.04^{***}	3 > 1 > 2
Openness	-.17 (1.05)	-.44 (1.21)	-.51 (1.02)	1.96	
High Intensity Pleasure	-.13 (.96)	1.26 (.76)	.37 (1.00)	45.62^{***}	2 > 3 > 1
Impulsivity	.90 (1.06)	2.61 (.84)	2.64 (.91)	81.52^{***}	1 < 2, 3
Inhibitory Control	-.98 (.69)	-2.33 (.83)	-2.39 (.86)	69.12^{***}	1 > 2, 3
Sadness	.10 (.75)	-.17 (.91)	1.44 (.89)	90.02^{***}	3 > 1, 2
Shyness	-.01 (1.10)	-.78 (.80)	.61 (1.09)	43.81^{***}	3 < 1 < 2
Soothability	-.30 (.68)	-.19 (.76)	-1.59 (.75)	100.83^{***}	3 < 1, 2
Assertiveness/Dominance	-.33 (1.14)	.68 (1.09)	.29 (1.11)	15.16^{***}	2 < 3 <1

Attentional Focus	-2.20 (.99)	-3.05 (.86)	-3.37 (.67)	39.12***	3 < 2 < 1
Low Intensity Pleasure	-.45 (.93)	-1.00 (1.22)	-.99 (1.11)	5.76**	1 > 2, 3
Perceptual Sensitivity	-.43 (.93)	-.33 (1.28)	-.17 (1.12)	1.12	

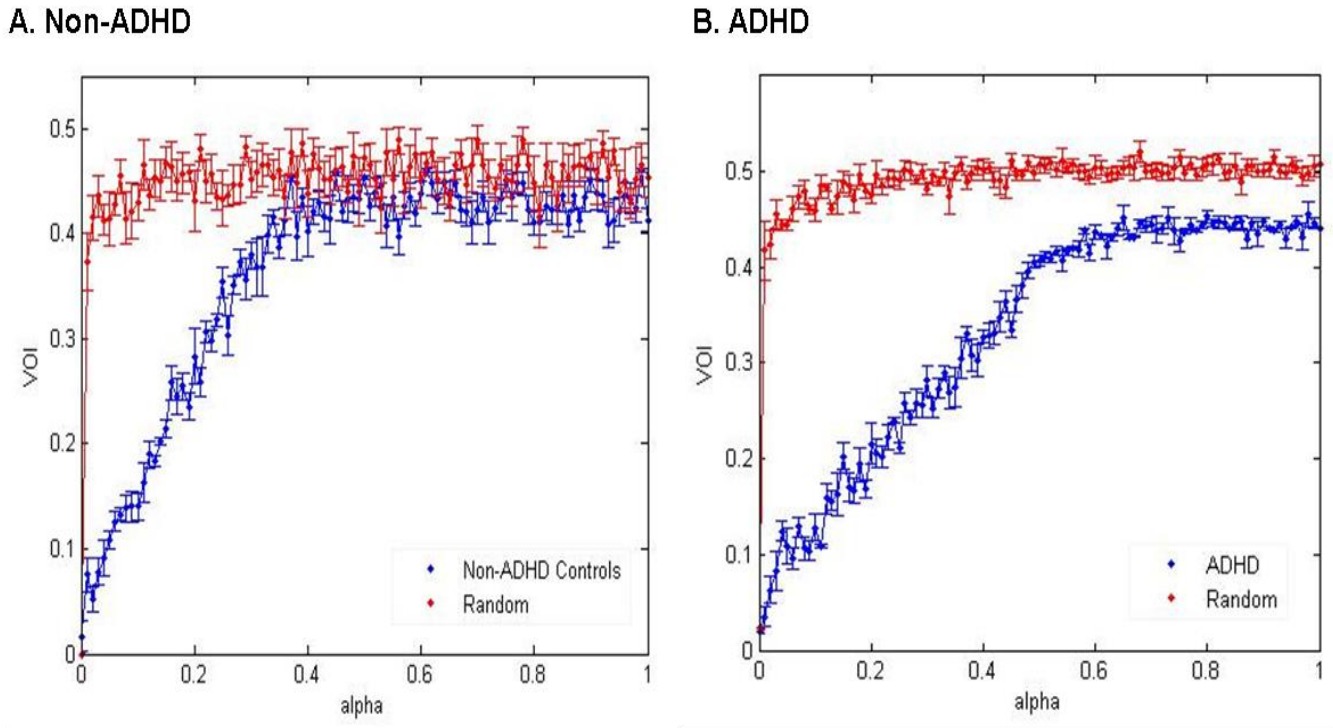
Note: I= ADHD Inattentive presentation; C= ADHD Combined presentation; SDQ= Strengths & Difficulties Questionnaire; ODD= Oppositional Defiant Disorder; T1= Time 1; T2= Time 2.

[a] Based on a subsample of 101 children with ADHD.

eTable 3. *Respiratory Sinus Arrhythmia (RSA; ms^2) and Pre-ejection Period (PEP; ms) in Emotion Induction/Suppression Task*

Variable	Control	Type 1 Mild	Type 2 Surgent	Type 3 Irritable	Post-hoc
	n=140	(n=48)	(n=60)	(n=60)	
Neutral picture baseline					
RSA	6.52 (.10)	7.06 (.17)	7.06 (.16)	6.97 (.16)	C < 1, 2, 3
PEP	97.94 (0.96)	100.74 (1.61)	94.23 (1.46)	99.03 (1.45)	2 < C, 1, 3
PEP task raw scores					
Negative induction	99.35 (1.04)	99.66 (1.70)	97.17 (1.57)	100.33 (1.51)	
Negative suppression	99.19 (1.02)	99.45 (1.67)	97.53 (1.53)	100.83 (1.47)	
Positive induction	98.51 (0.99)	100.30 (1.32)	94.94 (1.50)	100.31 (1.44)	
Positive suppression	99.65 (1.04)	102.42 (1.69)	97.00 (1.56)	99.60 (1.50)	
Grand Mean	99.22 (1.03)	100.29 (1.62)	96.76 (1.55)	100.26 (1.49)	
RSA task change scores					
Negative induction	0.42 (.07)	0.41 (.11)	0.18 (.11)	0.17 (.11)	
Negative suppression	0.45 (.07)	0.42 (.11)	0.18 (.11)	0.27 (.11)	
Positive induction	0.10 (.07)	0.29 (.12)	-0.14 (.11)	-0.08 (.11)	
Positive suppression	0.27 (.07)	0.34 (.11)	0.29 (.11)	0.29 (.11)	
Grand Mean	0.33 (.07)	0.38 (.11)	0.14 (.11)	0.16 (.11)	

eFigure. *Variation of Information (VOI)*



VOI analyses were used to assess group robustness. The red lines provide examples of groups detected due to random connections in the matrix. In this case, small changes in connections in the matrix (x-axis) result in large changes in group structure (y-axis). Blue lines indicate the groups detected for the non-ADHD control sample (**A**) and ADHD sample (**B**). In the case of the ADHD sample, in particular, small changes in connections result in similarly small change sin group structure indicating robust group assignments.

