Data Supplement for Ball et al., Multimodal Structural Neuroimaging Markers of Brain Development and ADHD Symptoms. Am J Psychiatry (doi: 10.1176/appi.ajp.2018.18010034)

Supplemental Methods

Participants

This paper reports on a subsample of the baseline cohort from the Children's Attention Project (CAP)(1). Participants were recruited in a two-stage screening and case-confirmation procedure using the Conners 3 ADHD Index (2) and a diagnostic interview (DISC-IV)(3).

As part of the 36-month follow-up, participants were invited to take part in a neuroimaging session and diagnostic status was re-assessed. Cognitive assessments and neuroimaging were acquired in 179 children (119 male; 83 ADHD; mean age at scan [range]=10.4 [9.4-11.9] years). After image acquisition and quality control, the final cohort comprised 160 individuals (104 male; 70 ADHD; mean age = 10.4 [9.7-11.9] years. The final group did not differ significantly from the full cohort in age (p=0.74), sex (p=0.77), socioeconomic status (p=0.94), diagnosis (p=0.56), or ADHD symptoms (hyperactive: p=0.77; inattentive: p=0.69). Twenty three individuals with ADHD were taking medication for their behavior (methylphenidate: n=21 {Concerta n=11, 1x 54mg, 9x 36mg, 1x 27mg, duration = 1month-4years [mean=20.6months]; Ritalin n=7, 10mg, duration = 10months – 4years [mean=31months]; Ritalin LA n=3, 2x 20mg [18months and 4years duration], 1x 30mg [4years]} and atomoxetine: n=2 {Strattera, 1x 18mg[2yrs 3months], 1x 25mg[6months]). At follow-up, 22 of the original ADHD group no longer met diagnostic criteria.

Clinical and cognitive assessment

Children took part in cognitive assessment, self-report survey and parent. With parental consent, questionnaires were also sent to the child's classroom teacher. Children completed the assessment under their usual classroom condition, therefore, if the child was currently using medication, they were not asked to cease medication for the assessment. Researchers were blind to diagnostic status and details of medication history and dosage were recorded at the end of the session.

Key assessment measures were broadly grouped into individual, clinical, cognitive, familial and perinatal factors:

Individual factors

Age, sex and self-reported handedness were recorded. Weight was measured using the average of two consecutive measurements using an Invicta portable stadiometer.

Pubertal stage was assessed using the Pubertal Development Scale (PDS). The primary caregiver was asked to rate their child's physical development on a four-point scale. This included questions assessing the presence of characteristics phenotypical of pubertal

onset such as deepening of voice and presence of facial hair in boys, and breast development and menarche for females. A combined PDS-Shirtcliff (PDSS) score was calculated.(4)

In addition, intracranial volume was estimated using the individual's T1-weighted image after MRI.

Clinical factors

ADHD symptom count (inattentive, and hyperactive-impulsive), and comorbid internalizing and externalizing behaviour were assessed using the DISC-IV. Children were classified as having an internalizing disorder if they met criteria for separation anxiety disorder, social phobia, generalized anxiety disorder, post-traumatic stress disorder, obsessive-compulsive disorder, hypomania or manic episode, and an externalizing disorder if they met criteria for oppositional defiant disorder or conduct disorder.

Additional measures of ADHD symptom severity were collected from the parent and teacher Conners 3 ADHD index. Autism spectrum symptoms were assessed using the Social Communication Questionnaire (SCQ-Lifetime version); a 40-item questionnaire measuring parent reported ASD symptoms.(5) Irritability was assessed using the parent-reported Affective Reactivity Index,(6) and social problems were assessed using the parent-reported subscale from the Strengths and Difficulties Questionnaire (SDQ).(7)

Details of any ADHD medication history and dosage were also recorded.

Cognitive factors

Baseline measures of intellectual function (IQ) was assessed at recruitment using the vocabulary and matrix reasoning subtests of the Wechsler Abbreviated Scale of Intelligence (WASI).

The matrix reasoning subtest of the WASI was repeated at follow-up as a measure of visuo-spatial reasoning. Academic cognitive functioning was assessed using the Word Reading and Math Computation subtests from the Wide Range Achievement Test 4 (WRAT 4),(8) the Clinical Evaluation of Language Fundamentals (fourth edition; CELF-4) screening test.(9) We also collected a teacher-reported 7 item measure of academic competence from the Social Skills Improvement System (SSIS).(10)

As part of a computerised battery, participants undertook several tasks, as follows: The Spatial N-Back: the 2-back version of this widely used measure of working memory requires flexible updating capabilities. Performance measure were *d'*, an index of the ability to discriminate between targets and non-targets, and *c*, which represents the participants' impulsivity to respond. The Stop Signal Task(11) assessed response inhibition; subjects perform a choice reaction task and, on a random selection of the trials,

an auditory stop signal instructed participants to withhold their response. The stop-signal reaction time (SSRT) assessed the ability to stop an initiated response. The Sustained Attention to Response Task (SART)(12) is a measure of sustained attention. The fixed version of SART is a repeating sequence of digits (1–9). Using a button press, participants respond to every digit (go-trial) except '3' (no-go trial). Performance measures of omissions, commissions and mean reaction time (ms) were used. Set Shifting: Two target pictures that vary along two dimensions (e.g., shape and color) were presented. Participants were cued with a letter to respond to the target pictures, according to one dimension, thus requiring cognitive flexibility. Performance was assessed using the switching cost (the differences between mean switch and non-switch trial reaction times within the heterogeneous block).

Familial factors

Neighborhood socioeconomic disadvantage (Socio-Economic Indexes for Areas Disadvantage Index (SEIFA)(13)) was recorded and parents reported on their years of schooling, number of children, and measures of parenting behaviors (warmth, angry, consistent) from the Longitudinal Study of Australian Children.(14)

Parenting behaviours were assessed measured using 6 items assessing parental warmth (e.g., how often in the past 6 months did you hug or hold this child for no apparent reason?), 5 items assessing angry parenting (e.g., how often are you angry when you punish this child?), and 6 items assessing consistent parenting (e.g., how often do you think that the level of punishment you give this child depends on your mood?) from LSAC.

To assess family quality of life, measures were derived using subscales from the Child Health Questionnaire (CHQ): family activities (e.g., how often has your child's behavior interrupted various everyday family activities (eating meals, watching TV); time impact (e.g., how often has your child's behavior caused you to cancel or change plans (personal or work) at the last minute; and emotional impact (e.g., how much worry or concern did child's emotional well-being or behavior cause you). Parental mental health was assessed using the Conners Adult ADHD Rating Scales (CAARS) for ADHD symptoms.(15) The Stressful Life Events Scale assessed a range of stressful life events experienced in the last 12 months.(16)

Perinatal factors

At the CAP baseline assessment (age 6-8 years), parents were asked to retrospectively recall risk factors including birth weight and whether time was required in the neonatal intensive care unit. Biological mothers were also asked to indicate whether they had drunk alcohol, or smoked cigarettes during pregnancy (0=abstained, 1=consumed during all three trimesters)

Image acquisition

All participants completed a 30 minute mock scanning session prior to the MRI scan in order to familiarise the child with the MRI procedure. Multimodal MRI data were acquired on a 3T Siemens Tim Trio MRI scanner (Erlangen, Germany) at the The Royal Children's Hospital, Melbourne. T1-weighted, multi-echo MPRAGE images were acquired with navigator-based prospective motion correction (MoCo) as follows: repetition time (TR) = 2530 ms; echo time (TE) = 1.77, 3.51, 5.32, 7.20 ms; inversion time (TI) = 1260 ms; flip angle = 7°, voxel size = 0.9mm3 isotropic. Diffusion weighted data were acquired using 25 diffusion gradient directions, b value = 1000 s/mm2, TR/TE = 3200/110 ms, voxel-size = 2.4mm3 isotropic, multiband factor = 3, with six interleaved b = 0 volumes. A reverse phase encoded image was acquired to correct for magnetic susceptibility-induced distortions during EPI acquisition.

Image pre-processing

Detailed quality control procedures were followed at each step, beginning with visual inspection of all volumes prior to analysis. Volumes with excess movement (n = 5), gross acquisition artefact (n = 1), or major anatomical abnormality (n = 2) were excluded.

Each participant's T1 image was intensity normalised, corrected for bias field inhomogeneities and aligned to MNI152 2mm space using diffeomorphic nonlinear registration in ANTs.(17,18) Voxel-wise maps of volume change induced by the transformation were characterised by the determinant of the Jacobian operator, referred to here as the Jacobian map.

Cortical thickness and surface ares were computed using FreeSurfer. This software performs brain extraction, intensity normalisation, and cortical segmentation and tessellation of the grey matter/white matter boundary followed by automated topology correction. Cortical geometry was matched across individual surfaces using spherical registration. Cortical surfaces were visually inspected by an experienced neuroimaging scientist (CM) and manual edits made where necessary, before regeneration of corrected cortical surfaces. To reduce computational load, surface data were downsampled from the *fsaverage* surface to the lower-resolution *fsaverage6* space (40962 vertices per hemisphere).

Prior to analysis, both tissue volume (Jacobian) maps and cortical thickness and area maps were smoothed with a Gaussian kernel of 10 mm FWHM.

Diffusion data were subjected to eddy current, motion, and susceptibility induced distortion correction using *topup* and *eddy* tools.(19) Images were skull-stripped, and diffusion tensors fit with weighted-least squares. Skeletonised FA and MD maps were created by projecting individual data onto a group-average mean FA skeleton (thresholded at FA ≥ 0.2 [default value]) in MNI152 1mm space.

Linked independent component analysis

We combined imaging modalities (tissue volume, cortical thickness, cortical area, FA and MD) using FSL's Linked ICA (FLICA) toolbox to extract a set of multimodal imaging features for statistical analysis.(20) We performed linked ICA on the full dataset (160 subjects × 5 modalities) for 10000 iterations. Following previous examples, and given our sample size, we initially specified the model to estimate 25 independent components,(21,22). This process yielded a set of multimodal components reflecting patterns of shared variance present across all image modalities (Figure S2). We also performed the decomposition specifying 20 and 30 components to assess model stability (Supplemental Results; Figure S4). Prior to linked ICA, variance normalisation was performed separately on each set of aligned image maps.(23) FLICA was implemented in Matlab R2015b (Natick, MA), code is available at https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLICA.

Canonical correlation analysis

Canonical correlation analysis (CCA) was performed using the *canoncorr* function in Matlab R2015b. Both imaging and phenotypic data were tested for normality before CCA. We used CCA to define multivariate associations between each subject's demographic and behavioural data and the multimodal imaging features derived from linked ICA (Figure S1A). CCA aims to identify linear relationships between datasets by maximising the correlation between two canonical variates, U and V, that are each constructed from a weighted sum of the variables in each dataset, X and Y, such that $U = aX$ and $V = bY$, where a and b are the canonical weights optimised by the model. Successive linear transforms are sought under the constraint that they are uncorrelated with other pairs. The strength of association between each pair of canonical variates is referred to as the canonical correlation.

CCA is sensitive to outliers and non-normally distributed data, we therefore applied a set of transformations to the phenotypic data to account for both non-normality and missing data in the original dataset. Following Smith et al.(24), data were first normalised by applying a rank-based inverse Gaussian transformation, before standardisation to zero mean and unit variance. We then performed a dimension reduction step using principal components analysis (PCA), retaining enough components to explain at least 90% of the variance in the original dataset. During PCA, we accounted for missing data by estimating the subject × subject covariance matrix in an elementwise fashion, calculating covariance for each pair of subjects while ignoring missing variables. This covariance matrix was then projected to the nearest valid symmetric positive definite matrix before eigenvalue decomposition. This process mitigates the risk of overfitting through dimension reduction, avoids the need for imputation of missing data values, and results in a set of normally distributed component weights for CCA.(24)

As some variables were missing for a relatively large number of participants (Supplementary Table 1), we also tested this approach after removing all variables with missing data for more than 20% of subjects (n=4), and all subjects missing more than 20% of variables (n=11), though we found that this did not significantly affect the model (Figure S4).

As both demographic and imaging components represented a linear combination of the original variables (derived from PCA and linked ICA, respectively), we can assess the contribution of each of the original variables to the final model by calculating the *loading*, or correlation, between the original variable (e.g.: $X_1, \ldots, X_n$) and their respective canonical variate ($U_1, \ldots, U_m$) in each pair, where n is the number of original variables in X and m is the number of canonical pairs.(24)

Validation cohort

T1-weighted images were downloaded and processed as in the NICAP cohort to derive tissue volume (Jacobian) maps in MNI152 space. Of 263 NYU cases, 17 failed initial QC inspection or were missing T1-weighted images, a further 7 were excluded due to a poor registration to the study template, and 8 were excluded due to no ADHD or hyperactivity data, resulting in n=231. Diffusion data were not available for this cohort but, as the spatial loading maps derived from linked ICA and CCA are inherently correlated across modalities, we are able to proceed using just tissue volume maps (thus also avoiding the need to also perform Freesurfer segmentations for each participant).

Reorganising the tissue volume maps into a 2D (voxel × subject) matrix, we performed a multivariate linear regression, regressing the spatial loading maps from the original analysis (as a voxel × component matrix) onto the NYU tissue volume data to estimate a set of weights that represent a subject-specific loading for each component map in the independent NYU cohort (Fig S1).(25)

Associations between component weight and age, ADHD diagnosis, degree of hyperactivity and inattentiveness measured using the CPRS, and IQ in the NYU cohort were then tested using linear regression in SPSS (v21; NY: IBM Corp.). To account for skew in some clinical measures, data were first normalised by applying an rank-based inverse Gaussian transformation before performing regression.

Head motion

Motion is an important consideration for neuroimaging studies of clinical populations.(26–28) Here, we estimate in-scanner motion using a measure of framewise displacement in both the NICAP and NYU cohorts (see Supplemental Results).

Supplemental Results

Model stability and motion considerations

Repeating the analysis after regressing ICV from the data; selecting a different number of linked ICA components to summarise the multimodal imaging data; removing variables or subjects with >20% missing data; and regressing out estimates of in-scanner motion. We found, for the most part, that the factor structures remained the same under these perturbations (Figure S4), with the exception of the fourth component, where some factor loadings (e.g. male sex, pubertal stage) varied across runs. This may be due to statistical power - as the canonical pairs are ordered according to strength, the fourth pattern represents the weakest of the significant correlations within the data – alternatively, the varying factor structures may reflect a different canonical pair was ordered fourth in these runs.

In-scanner motion can adversely affect metrics derived from both structural and diffusion-weighted MRI.(26–28) We performed a number of additional analyses in order to determine the impact of in-scanner motion on our observations. For each participant in the NICAP cohort, we calculated framewise displacement (FD) as the mean displacement between adjacent diffusion volumes as a marker of in-scanner motion.

In the NICAP cohort, FD did correlate with the number of hyperactive (R^2=0.04, p=0.44), or inattentive (R^2=0.001, p=0.725) symptoms. Regressing FD from the imaging and phenotypic data prior to CCA did not affect the number of significant canonical correlations, the phenotypic factor structure (Figure S4), or the associated spatial loading maps.

Using FD values calculated in the NYU cohort,(27) we found that in-scanner motion correlated with age (R^2=0.029, p=0.011), hyperactivity (R^2=0.03, p=0.007) and, to a lesser extent, inattentiveness (R^2=0.017, p=0.052). However, FD was not associated with expression of either imaging component 2 (p=0.19) or 3 (p=0.63), and the inclusion of FD as a covariate did not affect the reported associations between component expression and age and hyperactivity (see main Results section).

Finally, we used partial least squares regression to derive a single motion-related imaging component (r=0.63, p<0.001) in the NICAP cohort and, using multivariate linear regression, we derived an expression weight for this imaging marker in the NYU cohort. This motion component did not correlate with either hyperactivity (R^2=0.003, p=0.403), or inattentiveness (R^2=0.01, p=0.124) in the NYU cohort.

Neuroimaging markers and ADHD diagnosis

Although the aim of the study was to examine symptoms dimensionally, we also performed a post-hoc analysis to explore how each imaging marker varied categorically as a function of ADHD diagnosis. In subjects with a persistent ADHD diagnosis (i.e.: participants who passed diagnostic criteria for ADHD at both the baseline and follow-up CAP sessions), we observed a small, statistically significant effect of diagnosis on expression of the third imaging marker (ADHD n=58; Control n=80; t_{136}=2.09, Cohen's d=0.36, p=0.039) only. Participants were further stratified into subtypes: inattentive (n=30); hyperactive (n=5); or combined (n=23). Imaging marker strength was highest in individuals with the combined type (mean score = 0.62), followed by hyperactive (0.27) and inattentive (-0.11), though the effect of subtype did not reach significance (ANOVA: F2,55=3.10, p=0.053).

As ADHD symptom scores and ratings were included in the construction of each imaging marker as part of the CCA analysis, it is perhaps not surpirising that the derived imaging pattern should differ between diagnostic groups and this analysis may be considered optimistically biased. We therefore replicated this effect in the independent NYU cohort using the imaging marker from the NICAP cohort. No phenotypic measures from the NYU cohort were used in the derivation of this imaging marker. We found that expression of the third imaging pattern was significantly higher in ADHD participants in the NYU cohort (t_{229}=3.37, Cohen's d = 0.45, p<0.001), although there was no significant main effect of subtype within the ADHD cohort (ANOVA: F2,129=1.4, p=0.325). The comparable diagnostic effects (Cohen's d = 0.36 in NICAP compared to 0.44 in the NYU cohort) which are in line with other estimates of effect size between diagnostic groups in ADHD.(29) No other markers differed between groups. We also did not find a significant interaction between imaging score and ADHD diagnosis when predicting hyperactivity (p=0.15) or inattentiveness (p=0.29), supporting the dimensional view of ADHD, as the relationship between the imaging marker and hyperactive behaviour spanned diagnostic groups, confirming previous reports of structural and functional neuroanatomical correlates of ADHD behaviours in healthy controls.(30,31)

Hierarchical linear regression in the validation dataset

We performed an additional hierarchical linear regression to further explore the relationship between the third imaging marker and hyperactivity in the NYU cohort by including sex, age and IQ in the first level and marker strength in the second. The full model explained almost 10% of variance in NYU hyperactivity score (R^2=0.097). The addition of component weight to the model induced a significant increase in model goodness-of-fit (R^2 change=0.029, $F_{1,217}$=6.89, p=0.009). The addition of framewise displacement (in-scanner motion) to the model did not alter this relationship (final model: R^2=0.12, p<0.001; R^2 change=0.04, $F_{1,206}$=8.45, p=0.004).

TABLE S1. NICAP cohort demographic data

Demographic factors	Source	Mean (range)	Missing (%)
Individual factors			
Age - years	P	10.4 (9.4 -11.9)	
Weight - kg	D	39.3 (20.4 - 82.2)	3 (1.9)
Pubertal stage	P	1.4 (1.0 - 4.0)	8 (.05)
Male sex n(%)	P	104 (65)	
Left handed n(%)	C	26 (16.3)	1 (0.6)
Intracranial volume - mL	D	1236.3 (954.3 - 1512.2)	
Clinical factors			
Hyperactive symptoms	P	2.4 (0.0 - 9.0)	
Inattentive symptoms	P	3.6 (0.0 - 9.0)	
Medication status n(%)	P	23 (14.4)	
Parent's ADHD index	P	5.8 (0.0-20.0)	5 (3.1)
Teacher's ADHD index	T	3.5 (0.0 - 20.0)	28 (17.5)
Externalising disorder n(%)	P	45 (28.1)	1 (0.6)
Internalising disorder n(%)	P	26 (16.3)	1 (0.6)
ASD symptoms	P	4.9 (0.0 - 23.0)	3 (1.9)
Irritability	P	3.3 (0.0 -12.0)	8 (5.0)
Social difficulty	P	2.0 (0.0 - 9.0)	5 (3.1)
Cognitive factors			
IQ	D	100.1 (69 - 139)	
CELF: language	D	20.1 (9.0 - 32.0)	3 (1.9)
WRAT: maths	D	95.5 (64.0 - 130.0)	1 (0.6)
WRAT: reading	D	102.0 (55.0 - 145.0)	1 (0.6)
WASI: v/s reasoning	D	22.0 (7.0 - 33.0)	1 (0.6)
Academic competence	T	98.3 (64.0 - 121.0)	28 (17.5)
N-back: *c*	D	0.17 (-1.04 - 1.13)	28 (17.5)
N-back: *d'*	D	2.95 (0.18 - 4.39)	28 (17.5)
Stop-signal: SSRT - ms	D	294.7 (102.3 - 572.7)	28 (17.5)
SART: commission	D	5.6 (0.0 - 18.0)	39 (24.4)
SART: omission	D	7.0 (0.0 - 26.0)	39 (24.4)
SART: RT - ms	D	477.1 (234.3 - 741.8)	39 (24.4)
SS: Mean cost	D	158.8 (-1461.2 - 877.0)	64 (40)
Familial factors			
Parental education: year 9 or less (%)	P	8 (5)	3 (1.9)
year 10/11 (%)	P	31 (19.4)	
year 12 or above (%)	P	118 (73.8)	
No. siblings	P	1.8 (0.0 - 6.0)	3 (1.9)
Angry parenting	P	2.0 (1.0 - 4.8)	10 (6.3)
Consistent parenting	P	4.1 (1.3 - 5.0)	11 (6.9)
Parenting warmth	P	4.2 (2.54 - 5.0)	10 (6.3)
Parent ADHD symptoms	P	49.3 (33.0 - 90.0)	10 (6.3)
QoL: emotional	P	69.6 (0.0 - 100.0)	8 (5.0)
QoL: family	P	78.5 (0.0 - 100.0)	8 (5.0)
QoL: time	P	85.7 (0.0 - 100.0)	9 (5.6)
Stressful events	P	1.0 (0.0 - 5.0)	9 (5.6)
SEIFA	P	1017.7 (936.0 - 1128.0)	2 (1.3)

(Continued)

(Table S1 Continued)

Perinatal factors			
Birth weight - kg	P	3.41 (0.35 - 4.80)	5 (3.1)
NICU n(%)	P	29 (18.1)	3 (1.9)
Maternal alcohol† n(%)	P	17 (10.6)	15 (9.4)
Maternal smoking† n(%)	P	18 (11.3)	15 (9.4)

† alcohol and smoking levels: 1=smoked or consumed alcohol though all three trimesters
Source: C = child; D = direct assessment, P = Parent, T = teacher
ASD= autism spectrum disorder; CELF= Clinical Evaluation of Language Fundamentals NICU= time in neonatal intensive care unit; QoL= Quality of Life; RT= response time; SART= Sustain Attention to Response Time task; SES = Socioeconomic status; SS= Set Shifting task; SSRT= stop signal response time; WASI = Wechsler Abbreviated Scales of Intelligence; WRAT= Wide Range Achievement Test.

TABLE S2. ADHD-200 NYU cohort demographic data

Demographic factors	Mean (range)	Missing (%)
Age - years	11.6 (7.2 - 18.0)	
Male n(%)	146 (63.1)	
Intracranial volume - mL	1319.3 (1014.8 - 1639.4)	
ADHD index	60.6 (40.0 - 99.0)	
Hyperactive/impulsvity index	59.2 (41.0 - 90.0)	
Inattentive index	60.4 (40.0 - 90.0)	
IQ (range)	107.8 (73.0 - 142.0)	10 (4.3)

TABLE S3. Full phenotypic variate structure

	Pair 1: Head size			Pair 2: Development			Pair 3: ADHD symptoms			Pair 4: Cognitive profile		
Demographic factors	**Loading [95% CI]**	**Var_{exp}**	**p**	**Loading [95% CI]**	**Var_{exp}**	**p**	**Loading [95% CI]**	**Var_{exp}**	**p**	**Loading [95% CI]**	**Var_{exp}**	**p**
Individual factors												
Age	0.11 [0.01, 0.23]	0.01	0.162	0.31 [0.20, 0.43]	0.07	<0.001*	0.33 [0.21, 0.45]	0.05	<0.001*	-0.14 [-0.29, 0.00]	0.01	0.073
Weight	0.04 [-0.06, 0.16]	0.00	0.584	0.88 [0.86, 0.93]	0.57	<0.001*	0.06 [-0.04, 0.17]	0.00	0.423	-0.04 [-0.16, 0.07]	0.00	0.623
Pubertal stage	-0.19 [-0.30, -0.10]	0.03	0.019	0.34 [0.24, 0.44]	0.09	<0.001*	-0.27 [-0.39, -0.16]	0.04	<0.001*	-0.25 [-0.38, -0.13]	0.03	0.002
Male sex	0.49 [0.42, 0.59]	0.22	<0.001*	-0.11 [-0.21, 0.01]	0.01	0.166	0.45 [0.37, 0.55]	0.10	<0.001*	0.35 [0.27, 0.47]	0.06	<0.001*
Left handed	-0.02 [-0.12, 0.08]	0.00	0.802	0.05 [-0.07, 0.16]	0.00	0.562	0.22 [0.12, 0.33]	0.03	0.005	-0.14 [-0.25, -0.03]	0.01	0.087
Intracranial volume	0.94 [0.94, 0.96]	0.82	<0.001*	-0.09 [-0.20, 0.02]	0.01	0.262	0.05 [-0.06, 0.14]	0.00	0.561	0.00 [-0.10, 0.12]	0.00	0.965
Clinical factors												
Hyperactive symptoms	0.05 [-0.05, 0.17]	0.00	0.504	-0.17 [-0.28, -0.05]	0.02	0.035	0.39 [0.28, 0.50]	0.08	<0.001*	-0.21 [-0.33, -0.10]	0.02	0.009
Inattentive symptoms	0.04 [-0.07, 0.15]	0.00	0.611	-0.09 [-0.20, 0.03]	0.01	0.279	0.19 [0.08, 0.30]	0.02	0.064	-0.06 [-0.18, 0.05]	0.00	0.417
Medication status	0.01 [-0.10, 0.13]	0.00	0.860	-0.33 [-0.44, -0.23]	0.08	<0.001*	0.29 [0.19, 0.38]	0.04	<0.001*	-0.08 [-0.20, 0.05]	0.00	0.341
Parent's ADHD index	-0.06 [-0.18, 0.04]	0.00	0.412	-0.12 [-0.24, 0.00]	0.01	0.138	0.29 [0.18, 0.40]	0.04	<0.001*	-0.05 [-0.18, 0.06]	0.00	0.513
Teacher's ADHD index	0.12 [-0.01, 0.24]	0.01	0.191	-0.02 [-0.17, 0.11]	0.00	0.810	0.19 [0.06, 0.33]	0.02	0.029	0.11 [-0.02, 0.25]	0.01	0.223
Externalising symptoms	-0.04 [-0.16, 0.07]	0.00	0.584	-0.14 [-0.27, -0.02]	0.01	0.078	0.16 [0.03, 0.27]	0.01	0.048	0.01 [-0.14, 0.13]	0.00	0.891
Internalising symptoms	0.04 [-0.07, 0.16]	0.00	0.594	0.05 [-0.07, 0.17]	0.00	0.525	-0.21 [-0.33, -0.10]	0.02	0.008	0.09 [-0.04, 0.22]	0.00	0.242
ASD symptoms	-0.02 [-0.14, 0.09]	0.00	0.770	-0.08 [-0.19, 0.04]	0.00	0.328	0.37 [0.27, 0.48]	0.07	<0.001*	0.12 [0.00, 0.24]	0.01	0.130
Irritability	-0.24 [-0.36, -0.13]	0.05	0.003	-0.12 [-0.25, 0.01]	0.01	0.144	0.17 [0.05, 0.29]	0.01	0.035	0.23 [0.11, 0.36]	0.02	0.005
Social difficulty	-0.15 [-0.27, -0.04]	0.02	0.065	0.01 [-0.13, 0.12]	0.00	0.949	0.23 [0.12, 0.36]	0.03	0.004	0.11 [-0.01, 0.24]	0.01	0.160
Cognitive factors												
IQ	0.21 [0.10, 0.32]	0.04	0.008	0.07 [-0.05, 0.19]	0.00	0.392	-0.23 [-0.35, -0.12]	0.03	0.003	-0.09 [-0.22, -0.03]	0.00	0.249
CELF: language	0.20 [0.10, 0.32]	0.04	0.011	-0.06 [-0.19, 0.06]	0.00	0.496	-0.06 [-0.18, 0.06]	0.00	0.476	-0.31 [-0.44, -0.21]	0.04	<0.001*
WRAT: maths	0.39 [0.29, 0.50]	0.14	<0.001*	-0.19 [-0.31, -0.07]	0.03	0.019	-0.20 [-0.31, -0.09]	0.02	0.013	-0.07 [-0.19, 0.06]	0.00	0.409
WRAT: reading	0.30 [0.20, 0.41]	0.08	<0.001*	0.01 [-0.10, 0.13]	0.00	0.866	0.04 [-0.08, 0.17]	0.00	0.593	-0.18 [-0.31, -0.05]	0.01	0.026
WASI: v/s reasoning	0.31 [0.20, 0.41]	0.09	<0.001*	0.01 [-0.13, 0.11]	0.00	0.937	-0.15 [-0.28, -0.04]	0.01	0.060	-0.23 [-0.36, -0.13]	0.02	0.003
Academic competence	0.13 [0.00, 0.26]	0.02	0.142	-0.17 [-0.30, -0.04]	0.02	0.059	-0.03 [-0.15, 0.10]	0.00	0.768	-0.30 [-0.44, -0.18]	0.04	<0.001*
N-back: c	0.05 [-0.08, 0.16]	0.00	0.594	-0.13 [-0.26, -0.02]	0.01	0.137	-0.02 [-0.16, 0.11]	0.00	0.789	0.18 [0.05, 0.33]	0.02	0.036
N-back: d'	0.13 [0.01, 0.26]	0.02	0.137	-0.25 [-0.39, -0.12]	0.04	0.005	-0.14 [-0.27, 0.00]	0.01	0.111	0.14 [0.00, 0.29]	0.01	0.103
Stop-signal: RT	-0.05 [-0.18, 0.07]	0.00	0.555	0.15 [0.03, 0.29]	0.02	0.079	0.28 [0.15, 0.40]	0.03	0.001	0.17 [-0.04, 0.31]	0.00	0.059
SART: commission	0.00 [-0.13, 0.12]	0.00	0.968	-0.08 [-0.22, 0.05]	0.00	0.381	-0.05 [-0.21, 0.08]	0.00	0.594	0.13 [-0.02, 0.29]	0.01	0.156
SART: omission	-0.15 [-0.29, -0.02]	0.02	0.111	0.04 [-0.12, 0.18]	0.00	0.697	-0.02 [-0.16, 0.13]	0.00	0.871	-0.06 [-0.22, 0.10]	0.00	0.531
SART: RT	-0.35 [-0.48, -0.24]	0.12	<0.001*	0.10 [-0.02, 0.23]	0.01	0.232	-0.24 [-0.38, -0.11]	0.03	0.009	0.01 [-0.15, 0.11]	0.00	0.889
SS: Mean cost	-0.27 [-0.43, -0.11]	0.07	0.008	-0.13 [-0.30, 0.03]	0.01	0.204	-0.16 [-0.30, -0.01]	0.01	0.130	-0.27 [-0.43, -0.12]	0.03	0.009

(Continued)

(Table S3 Continued)

Familial factors												
Parental education	0.05 [-0.06, 0.16]	0.00	0.514	-0.08 [-0.20, 0.03]	0.00	0.322	-0.23 [-0.35, -0.12]	0.03	0.003	-0.17 [-0.29, -0.05]	0.01	0.036
No. siblings	-0.19 [-0.30, -0.09]	0.03	0.015	-0.24 [-0.35, -0.13]	0.04	0.004	0.06 [-0.06, 0.16]	0.00	0.491	-0.07 [-0.20, 0.04]	0.00	0.343
Angry parenting	-0.06 [-0.19, 0.06]	0.00	0.443	-0.12 [-0.25, 0.01]	0.01	0.147	0.25 [0.14, 0.38]	0.03	0.002	0.05 [-0.08, 0.18]	0.00	0.581
Consistent parenting	0.12 [0.02, 0.24]	0.01	0.134	-0.10 [-0.21, 0.01]	0.01	0.238	-0.31 [-0.44, -0.19]	0.05	<0.001*	0.10 [-0.04, 0.23]	0.00	0.236
Parenting warmth	-0.05 [-0.15, 0.06]	0.00	0.569	0.00 [-0.12, 0.12]	0.00	0.998	0.20 [0.10, 0.32]	0.02	0.012	0.10 [0.03, 0.22]	0.00	0.234
Parent ADHD symptoms	-0.10 [-0.22, 0.01]	0.01	0.236	0.04 [-0.08, 0.16]	0.00	0.654	0.04 [-0.08, 0.18]	0.00	0.586	0.09 [-0.05, 0.22]	0.00	0.262
QoL: emotional	0.15 [0.05, 0.27]	0.02	0.058	0.17 [0.04, 0.28]	0.02	0.041	-0.21 [-0.33, -0.08]	0.02	0.011	-0.13 [-0.26, -0.01]	0.01	0.103
QoL: family	0.02 [-0.09, 0.14]	0.00	0.820	0.14 [0.03, 0.28]	0.01	0.079	-0.29 [-0.40, -0.18]	0.04	<0.001*	-0.07 [-0.19, 0.05]	0.00	0.401
QoL: time	0.13 [0.03, 0.25]	0.02	0.101	0.15 [0.02, 0.27]	0.02	0.074	-0.27 [-0.39, -0.16]	0.04	<0.001*	-0.14 [-0.26, -0.02]	0.01	0.010
Stressful events	-0.04 [-0.15, 0.06]	0.00	0.593	0.03 [-0.07, 0.14]	0.00	0.687	0.38 [0.27, 0.49]	0.07	<0.001*	-0.02 [-0.15, 0.12]	0.00	0.800
SEIFA	0.22 [0.12, 0.32]	0.04	0.006	-0.16 [-0.29, -0.05]	0.02	0.041	-0.09 [-0.21, 0.03]	0.00	0.280	0.08 [-0.03, 0.22]	0.00	0.320
Perinatal factors												
Birth weight	0.22 [0.11, 0.32]	0.04	0.007	0.08 [-0.03, 0.20]	0.00	0.334	-0.01 [-0.13, 0.11]	0.00	0.919	0.07 [-0.05, 0.21]	0.00	0.359
NICU	0.10 [0.00, 0.20]	0.01	0.221	-0.11 [-0.22, 0.00]	0.01	0.189	0.26 [0.17, 0.38]	0.04	<0.001*	0.06 [-0.06, 0.18]	0.00	0.485
Maternal alcohol	-0.09 [-0.20, 0.02]	0.01	0.301	0.16 [0.05, 0.26]	0.02	0.060	0.18 [0.08, 0.28]	0.02	0.029	0.10 [-0.03, 0.23]	0.00	0.218
Maternal smoking	-0.08 [-0.20, -0.03]	0.01	0.330	-0.08 [-0.19, 0.05]	0.00	0.367	0.02 [-0.14, 0.17]	0.00	0.783	0.43 [0.35, 0.54]	0.08	<0.001*

□variance explained by respective imaging marker *loading significant at p<0.05, after Bonferroni correction for multiple comparisons across demographic factors. ASD= autism spectrum disorder; CELF= Clinical Evaluation of Language Fundamentals NICU= time in neonatal intensive care unit; QoL= Quality of Life; RT= response time; SART= Sustain Attention to Response Time task; SEIFA = Socio-Economic Indexes for Areas; SS= Set Shifting task; WASI = Wechsler Abbreviated Scales of Intelligence; WRAT= Wide Range Achievement Test.

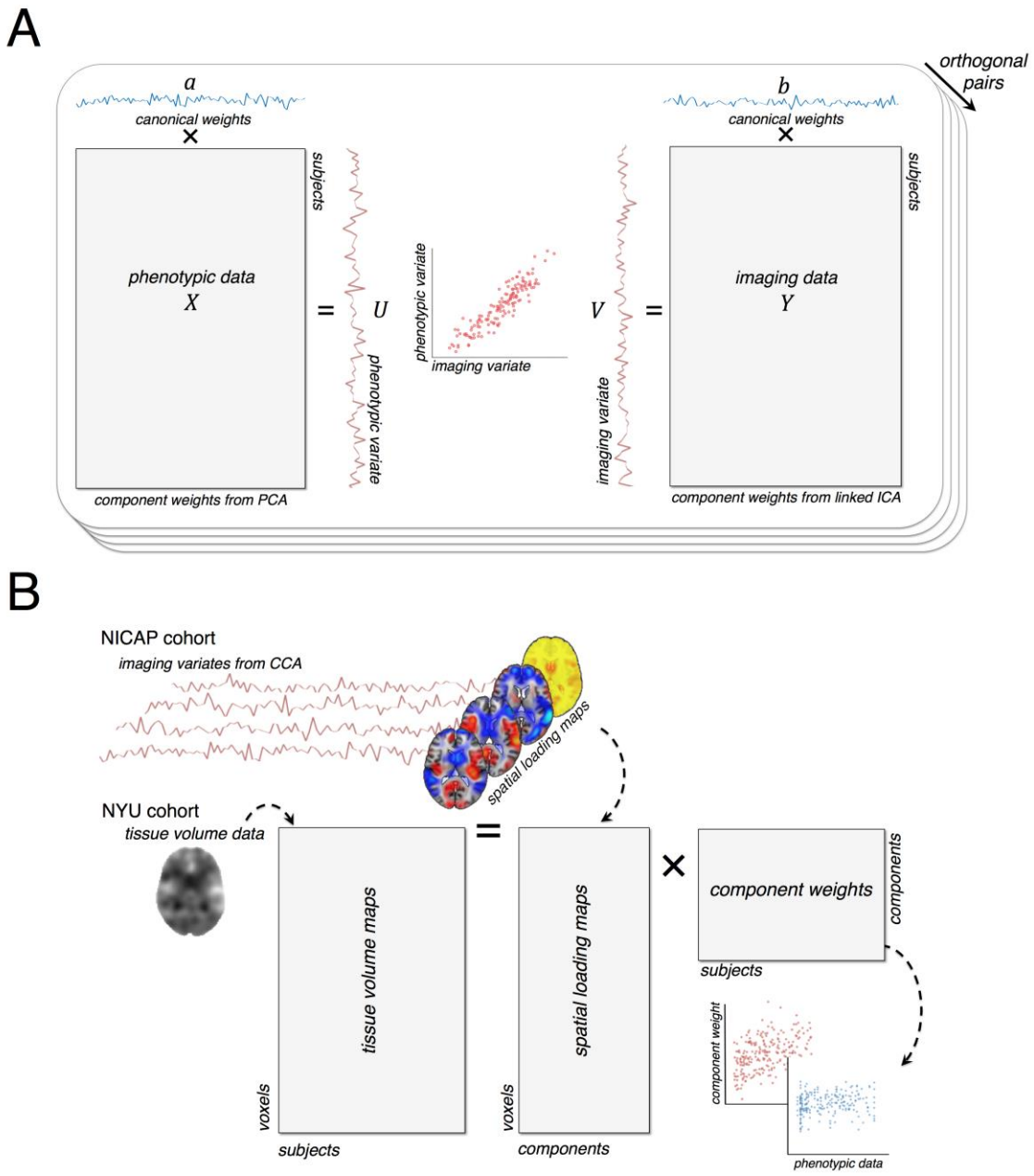

FIGURE S1. Analysis pipeline. Canonical correlation analysis (CCA) is illustrated in A. Multivariate associations between phenotypic and imaging data (after dimension reduction with principal component analysis (PCA) and linked independent component analysis (ICA), respectively) are sought by calculating model weights, *a* and *b*, that maximize the correlation between the phenotypic and imaging variates, *U* and *V*. Phenotypic loadings and spatial loading maps are calculated by correlating the canonical variates with the original phenotypic variables and the original imaging data. In B, subject-specific component weights are estimated for an independent cohort using spatial loading maps derived from the original CCA using a multivariate spatial regression. Associations are then sought between estimated component weights and phenotypic data.

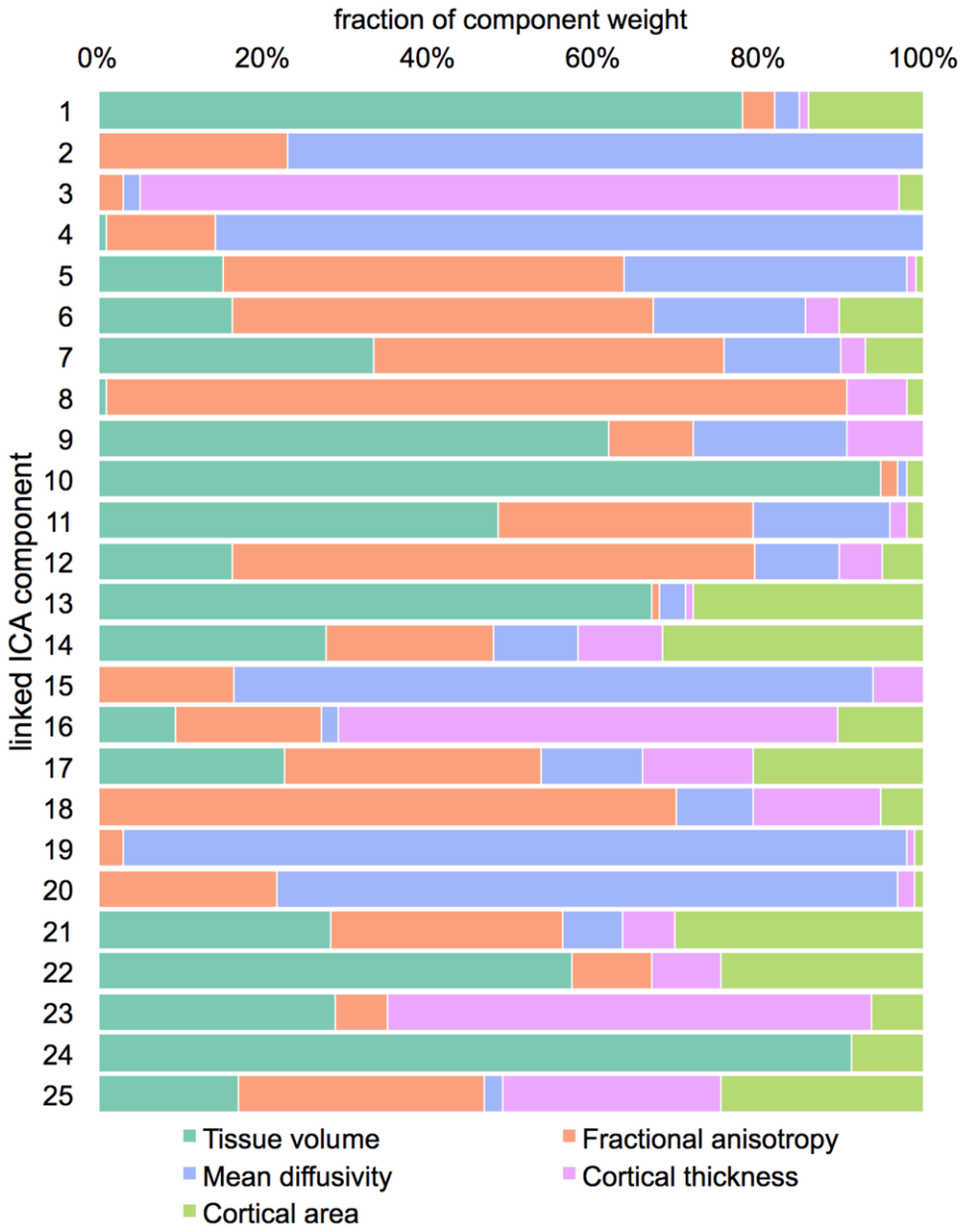

FIGURE S2. Contribution of each modality to linked ICA The relative weight of each modality across 25 linked ICA components. All components represented shared patterns of variance across all modalities.

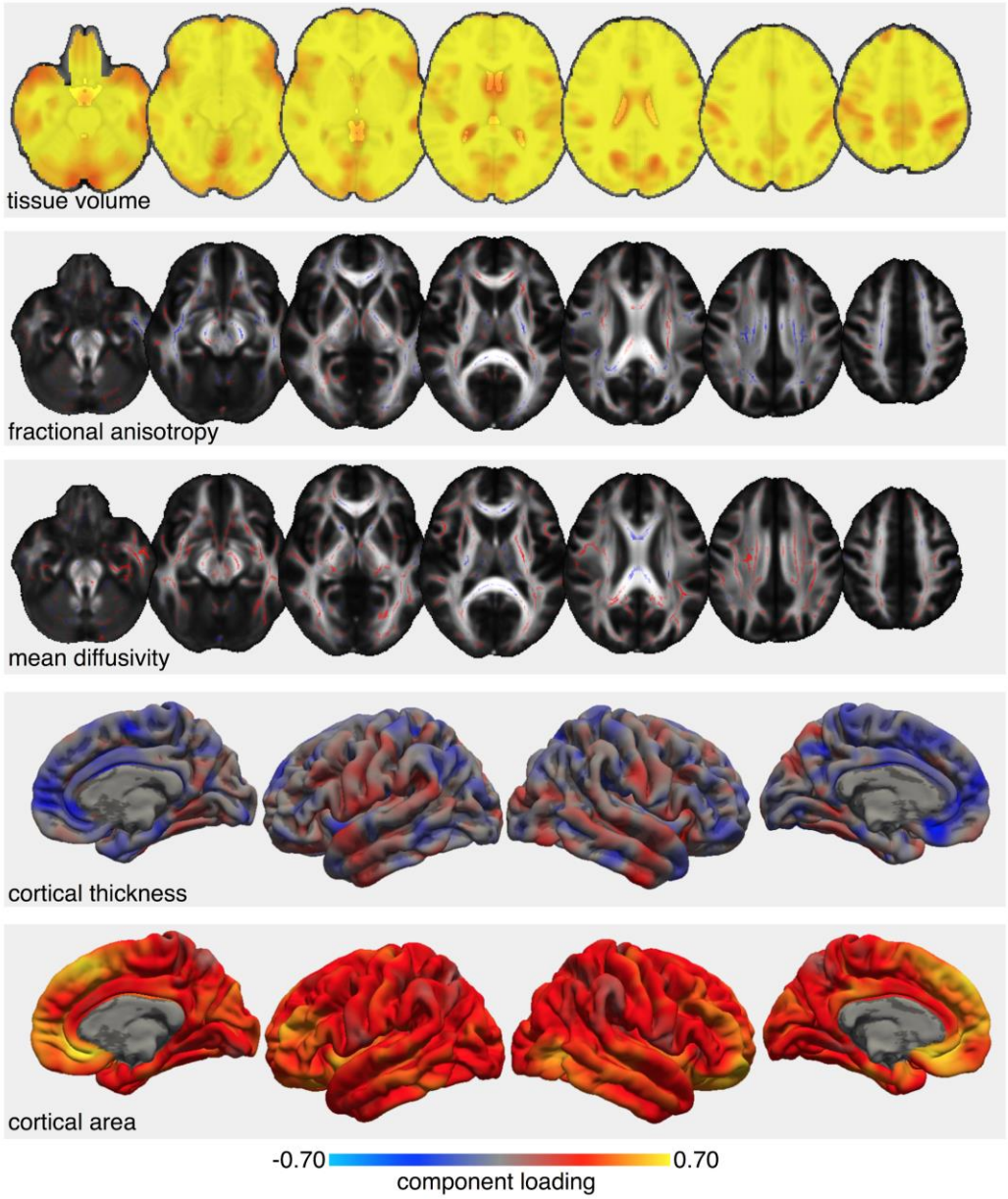

FIGURE S3. Multimodal imaging marker of head size. Voxelwise correlations between the first canonical imaging variate and each imaging modality are shown. The strength of correlation is shown by the colourbar. The corresponding clinical factor structure is shown in Supplementary Table 3. Images are available to view on Neurovault (http://neurovault.org/collections/2277/).

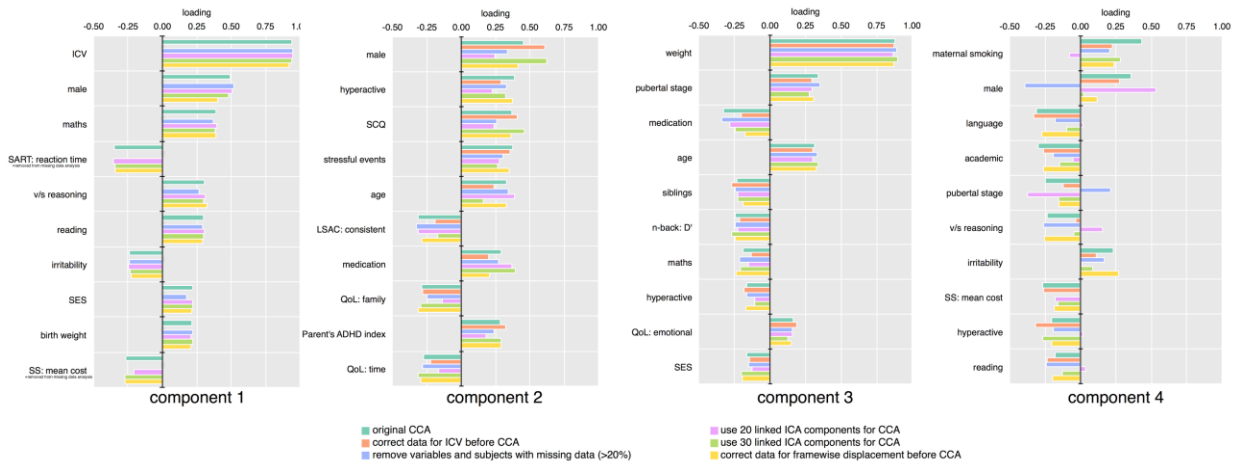

FIGURE S4. Stability of phenotypic variate structure over different parameter settings. The ten phenotypic variables with the largest loading on each of the four phenotypic factors in the original CCA are shown. Factor loadings are shown for each after performing CCA with several alternative parameter settings: regressing ICV or framewise displacement from the data prior to CCA, removing variables (n=4) or subjects (n=11) with a large amount of missing data, and using 20 or 30 linked ICA components as imaging variables.

Additional References

1. Sciberras E, Efron D, Schilpzand EJ, Anderson V, Jongeling B, Hazell P, et al. The Children's Attention Project: a community-based longitudinal study of children with ADHD and non-ADHD controls. BMC Psychiatry. 2013 Jan 10;13:18.

2. Conners CK. Conners 3rd Edition (Conners 3). 3rd ed. Toronto: Multi-Health Systems; 2008.

3. Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. 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. 2000 Jan;39(1):28–38.

4. Shirtcliff EA, Dahl RE, Pollak SD. Pubertal Development: Correspondence between hormonal and physical development. Child Dev. 2009;80(2):327–37.

5. Rutter M, Bailey A, Lord C. The Social Communication Questionnaire. Los Angeles, CA: Western Psychological Services; 2003.

6. Stringaris A, Goodman R, Ferdinando S, Razdan V, Muhrer E, Leibenluft E, et al. The Affective Reactivity Index: a concise irritability scale for clinical and research settings. J Child Psychol Psychiatry. 2012 Nov;53(11):1109–17.

7. Hawes DJ, Dadds MR. Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. Aust N Z J Psychiatry. 2004 Aug;38(8):644–51.

8. Wilkinson G, Robertson G. Wide Range Achievement Test (WRAT4). Lutz: Psychological Assessment Resources; 2006.

9. Semel E, Wiig E, Secord W. Clinical Evaluation of Language Fundamentals-Screening Test Australian & New Zealand Language Adapted Edition (CELF-4 Screener). Sydney: NCS Pearson; 2004.

10. Gresham FM, Elliott SN, Kettler RJ. Base rates of social skills acquisition/performance deficits, strengths, and problem behaviors: an analysis of the Social Skills Improvement System--Rating Scales. Psychol Assess. 2010 Dec;22(4):809–15.

11. Verbruggen F, Logan GD, Stevens MA. STOP-IT: Windows executable software for the stop-signal paradigm. Behav Res Methods. 2008 May;40(2):479–83.

12. Robertson IH, Manly T, Andrade J, Baddeley BT, Yiend J. "Oops!": performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia. 1997 Jun;35(6):747–58.

13. Australian Bureau of Statistics. Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia. Australian Bureau of Statistics; 2011.

14. Zubrick SR, Smith GJ, Nicholson J, Sanson A, Jackiewicz TA. Parenting and Families in Australia [Internet]. Rochester, NY: Social Science Research Network; 2008 Jun [cited 2017 Aug 21]. Report No.: ID 1703269. Available from: https://papers.ssrn.com/abstract=1703269

15. Conners K, Erhardt D, Sparrow E. Conners Adult ADHD Rating Scales (CAARS). Toronto: Multi-Health Systems; 1999.

16. Brugha TS, Cragg D. The List of Threatening Experiences: the reliability and validity of a brief life events questionnaire. Acta Psychiatr Scand. 1990 Jul;82(1):77–81.

17. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal. 2008 Feb;12(1):26–41.

18. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: Improved N3 Bias Correction. IEEE Trans Med Imaging. 2010 Jun;29(6):1310–20.

19. Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. NeuroImage. 2016 Jan 15;125:1063–78.

20. Groves AR, Beckmann CF, Smith SM, Woolrich MW. Linked independent component analysis for multimodal data fusion. NeuroImage. 2011 Feb 1;54(3):2198–217.

21. Groves AR, Smith SM, Fjell AM, Tamnes CK, Walhovd KB, Douaud G, et al. Benefits of multi-modal fusion analysis on a large-scale dataset: life-span patterns of inter-subject variability in cortical morphometry and white matter microstructure. NeuroImage. 2012 Oct 15;63(1):365–80.

22. Francx W, Llera A, Mennes M, Zwiers MP, Faraone SV, Oosterlaan J, et al. Integrated analysis of gray and white matter alterations in attention-deficit/hyperactivity disorder. NeuroImage Clin [Internet]. 2016 [cited 2016 Mar 15]; Available from: http://www.sciencedirect.com/science/article/pii/S2213158216300468

23. Beckmann CF, Smith SM. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans Med Imaging. 2004 Feb;23(2):137–52.

24. Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TEJ, Glasser MF, et al. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat Neurosci. 2015 Nov;18(11):1565–7.

25. Nickerson LD, Smith SM, Öngür D, Beckmann CF. Using Dual Regression to Investigate Network Shape and Amplitude in Functional Connectivity Analyses. Front

Neurosci [Internet]. 2017 Mar 13 [cited 2017 Aug 21];11. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5346569/

26. Alexander-Bloch A, Clasen L, Stockman M, Ronan L, Lalonde F, Giedd J, et al. Subtle In-Scanner Motion Biases Automated Measurement of Brain Anatomy From In Vivo MRI. Hum Brain Mapp. 2016 Jul;37(7):2385–97.

27. Pardoe HR, Kucharsky Hiess R, Kuzniecky R. Motion and morphometry in clinical and nonclinical populations. NeuroImage. 2016 Jul 15;135:177–85.

28. Yendiki A, Koldewyn K, Kakunoori S, Kanwisher N, Fischl B. Spurious group differences due to head motion in a diffusion MRI study. NeuroImage. 2014 Mar;88:79–90.

29. Hoogman M, Bralten J, Hibar DP, Mennes M, Zwiers MP, Schweren LSJ, et al. Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. Lancet Psychiatry. 2017 Apr;4(4):310–9.

30. Shaw P, Gilliam M, Liverpool M, Weddle C, Malek M, Sharp W, et al. Cortical development in typically developing children with symptoms of hyperactivity and impulsivity: support for a dimensional view of attention deficit hyperactivity disorder. Am J Psychiatry. 2011 Feb;168(2):143–51.

31. Rosenberg MD, Finn ES, Scheinost D, Papademetris X, Shen X, Constable RT, et al. A neuromarker of sustained attention from whole-brain functional connectivity. Nat Neurosci. 2016 Jan;19(1):165–71.