bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Data-driven subtyping of behavioural problems associated with ADHD in children

Joe Bathelt, Joni Holmes, the CALM team & Duncan E. Astle

MRC Cognition and Brain Sciences Unit

University of Cambridge

15 Chaucer Road

Cambridge CB2 7EF

Email: joe.bathelt@mrc-cbu.cam.ac.uk

Abstract

Many developmental disorders are associated with deficits in controlling and regulating behaviour. These difficulties are foremost associated with attention deficit hyperactivity disorder (ADHD), but are also frequently observed in other groups, including children with diagnoses of specific learning difficulties, autistic spectrum disorder, or conduct disorder. The co-occurrence of these behavioural problems across disorders typically leads to comorbid diagnoses and can complicate intervention approaches. An alternative to classifying children on the basis of specific diagnostic criteria is to use a data-driven grouping that identifies dimensions of behaviour that meaningfully distinguish groups of children and become specific targets for intervention. The current study applies a novel data-driven clustering algorithm to group children by similarities in their ratings on a parent questionnaire that is commonly used to assess behavioural problems associated with ADHD. The sample consisted of 442 children identified by health and educational professionals as having difficulties in attention, learning and/or memory. The data-driven clustering yielded three distinct groups of children with symptoms of either: (1) elevated inattention and hyperactivity/impulsivity, and poor executive function, (2) learning problems, and (3) aggressive behaviour and problems with peer relationships. These groups were associated with significant inter-individual variation in white matter connectivity of the prefrontal and anterior cingulate. In sum, data-driven classification of executive function difficulties identifies stable groups of children, provides a good account of inter-individual differences, and aligns closely with underlying neurobiological substrates.

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

ADHD is commonly diagnosed on the basis of elevated symptoms of inattention and hyperactivity/impulsivity (ICD-10 F90.9, World Health Organization, 1992). These problems are typically linked to difficulties with executive function (Barkley, 1997; Castellanos & Tannock, 2002; Pennington & Ozonoff, 1996) and aggressive behaviours (Harty, Miller, Newcorn, & Halperin, 2008), and can lead to learning difficulties (Czamara et al., 2013) and problems making and sustaining friendships (Mrug et al., 2012). Symptoms of inattention and hyperactivity/impulsivity are also common among other developmental disorders including dyslexia (Germano, Gagliano, & Curatolo, 2004), autism (Corbett, Constantine, Hendren, Rocke, & Ozonoff, 2009) and conduct disorder (Szatmari, Boyle, & Offord, 1989). This symptom overlap complicates research into the causes and treatment of common developmental difficulties. In practical terms, this often results in children receiving comorbid diagnoses and standard intervention approaches for particular disorders (e.g. medication for ADHD) that often do not meet their individual needs. The aim of the current study was to use a data-driven approach to identify groups of children with similar dimensions of behavioural problems and to investigate the relationship between white matter connectivity and these groupings.

Traditional categorical diagnostic approaches to understanding ADHD and other developmental disorders (e.g. ICD-10, World Health Organization, 1992) have considerable practical advantages by facilitating clinical decision making (Sonuga-Burke, 1998). However, current diagnoses are based on a tradition of clinical insight rather than knowledge about pathophysiological mechanisms (Sonuga-Burke & Halperin, 2010). The slow progress in understanding the mechanisms leading to neurodevelopmental disorders, and associated difficulties with identifying effective treatments, may in large part be attributable to the heterogeneous aetiology within traditional diagnostic classes. Subgroups within diagnostic categories may follow different pathways with a superficially similar clinical presentation (Nigg, 2006).

An alternative to classifying children on the basis of specific diagnostic criteria is to use a data-driven grouping that identifies the dimensions of behaviour that meaningfully distinguish groups of children. Recent evidence from taxonomic and heritability studies support this approach to understanding ADHD (Marcus & Barry, 2011, Gjone, Stevenson, & Sundet, 1996). A dimensional approach can provide the practical advantages of clearly

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

defined categories, while also identifying the most pertinent behavioural characteristics for investigations into pathophysiological mechanisms. The current study uses a data-driven community clustering approach to do just this. This alternative approach is made possible by recent advances in network science methods, which we apply to a large dataset comprised of children identified as having problems in attention, learning and/or memory, who were referred to the study by educational and clinical professionals working in various specialist children's services. This large sample includes children with specific, multiple and no diagnoses, and is not already restricted to children who have met particular diagnostic criteria. This offers the opportunity to identify the most pertinent behavioural dimensions, while side-stepping the biases inherent in recruiting according to current diagnostic classification.

Data-driven identification of behavioural profiles is not a straightforward task. Most clustering algorithms necessitate *a priori* assumptions, like the geometrical properties of the cluster shape, the tuning of some parameters, or setting the number of desired clusters. These assumptions are difficult to make with psychometric data, but network science provides a possible solution. Network science is the study of complex networks, which represent relationships between data as a network of nodes connected by edges. This methodological approach provides mathematical tools for quantifying the organisation of networks and the relationships between the nodes within them (Bullmore & Sporns, 2009). Defining subdivisions of highly-connected nodes within a network, so called communities, is an area of network science that has received considerable attention as it applies to many real world problems (Barabasi, 2016). In the case of psychometric data, the network can represent the similarity of scores between participants. Community detection makes it possible to define subgroups of participants that are most similar, while being as distinct as possible from other subgroups. The aim of the current study was to investigate whether subtypes of behavioural problems commonly linked with ADHD can be identified in a large sample of children referred for cognitive and learning difficulties using a community detection approach based on parent ratings on the Conner's questionnaire. This scale is routinely administered in health care and educational settings, and in many clinics in the U.K. it is used to measure behavioural problems at home and school to aid in the diagnosis of ADHD.

One of the aims of data-driven nosology is to identify behavioural dimensions that are more closely related to biological mechanisms. In the current study, we explored differences in white-matter connectivity between groups identified through community detection. White matter maturation is a crucial process of brain development that extends into the third decade of life (Lebel, Walker, Leemans, Phillips, & Beaulieu, 2008), which has been found to relate closely to cognitive development (Bathelt, Gathercole, Johnson, & Astle, *in press*; Clayden et al., 2011; Stevens, Skudlarski, Pearlson, & Calhoun, 2009). White matter maturation is thought to support cognitive development through better communication and integration between brain regions, particularly over longer distances (Collin & van den Heuvel, 2013). Accordingly, the brain can be modelled as a network of brain regions connected by white matter, commonly referred to as a connectome. Brain regions vary in the number of their connections – their node degree - which gives an indication of their importance for the network (Rubinov & Sporns, 2010; van den Heuvel,

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Kahn, Goni, & Sporns, 2012). To explore which brain regions were most closely linked to the behavioural profiles identified through consensus clustering, we used a multi-variate dimension-reduction technique called partial least squares (PLS) (Krishnan, Williams, McIntosh, Abdi, 2011). In our analysis, PLS defined brain components that maximally distinguished the behaviourally defined groups.

Participants and Methods

Participants

The sample consisted of ratings on 442 children, using the Conner's Parent Rating Short Form 3rd edition (Conners, 2013), referred to as "Conners 3" from here on. The ratings were completed by parents or caregivers (age: mean=110.51 months; SE=1.24; Range=62-215; 295 male) as part of a larger ongoing study at the Centre for Attention, Learning and Memory (CALM) at the MRC Cognition and Brain Sciences Unit. Children were recruited to the CALM research clinic on the basis of having problems in attention, learning and memory that had come to the attention of a professional working in schools (e.g. special needs teacher) or specialist children's community services (e.g. clinical or educational psychologists, speech and language therapists or paediatricians). During the clinic visit, children completed a wide range of cognitive assessments while their parents/caregivers filled in questionnaires about the child's behaviour. Children were also invited for an MRI structural scan (see Figure 1 for attainment). The data reported here include three questionnaires and the MRI data. Exclusion criteria for referrals were significant or severe known neurological disorders, problems in vision or hearing that were uncorrected, or having a native language other than English. This study was carried out in accordance with the Declaration of Helsinki and was approved by the local NHS research ethics committee (Reference: 13/EE/0157). Written parental/caregiver consent was obtained and children provided verbal assent.

Some children in the broad sample of children referred for problems relating to attention, learning, and/or memory, had received diagnoses through standard community services (see Table 1 for a breakdown of diagnoses). Among the children with a diagnosis, ADHD was the most common. Other diagnostic labels were rare. Therefore, diagnostic labels were grouped together for the comparison between diagnoses and the data-driven groups. Primary diagnoses of dyslexia, dyscalculia, or dysgraphia were summarised as 'learning deficits'. Primary diagnoses of autism spectrum disorder, autism, or Asperger syndrome were summarised as 'ASD'. Other labels, like OCD, depression, anxiety, or developmental delay occurred only in a few individuals and were grouped as 'other'.

Table 1: Breakdown of children by pre-existing diagnoses and referral routes. Abbreviations: ADHD: attention deficit hyperactivity disorder, ASD: autism spectrum disorder, SENCo: special educational needs coordinator

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Diagnosis	Total	%
None	302	76.7
ADHD	61	15.6
Learning Deficit	32	8.2
ASD	24	6.2
Other	23	5.9

Referrer	Total	%
SENCo	262	66.9
Paediatrician	82	21.0
Clinical Psychologist	29	7.4
Speech and Language Therapist	29	7.4
Specialist Teacher	13	3.3
ADHD Nurse Practitioner	13	3.3
Educational Psychologist	6	1.5
Family worker locality team	5	1.3
Child Psychiatrist	2	0.5
Private tutor	1	0.3

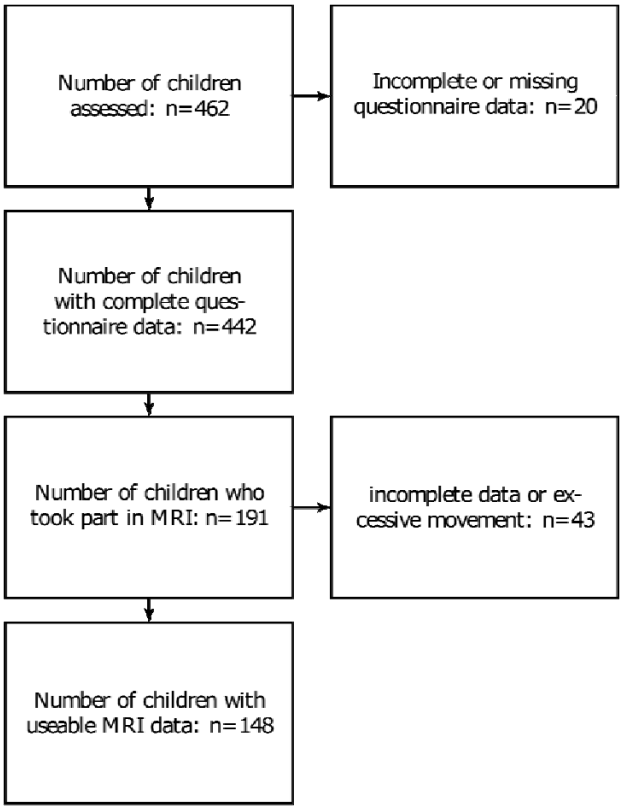


Figure 1: Overview of data included in behavioural and connectome analysis

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Behavioural Analysis

Questionnaire Data

We used the Conners-3 questionnaire for parents (Conners, 2013), which is designed to assess behavioural difficulties associated with ADHD and related disorders. This scale is well validated with good psychometric integrity (Internal consistency: Cronbach's alpha=0.91 [Range: 0.85-0.94]; Factorial validity: RMSEA=0.07 based on confirmatory factor analysis in a replication sample; for details see Conners, 2013). Questionnaire items are summarised into six scales (Inattention, Hyperactivity/Inattention, Learning Problems, Executive Function, Aggression, Peer Problems) and a total ADHD score is also derived. T scores of 60 and above are indicative of clinical levels. A high proportion of children in the sample had scores in this range on each scale of the subscales (see Table 2).

Table 2: Scores on each scale of the Conners-3 questionnaire (inattention, hyperactivity/impulsivity, learning problems, executive function, aggression, peer relationships) for the entire sample. The last two columns indicate the total number and the percentage of children in the sample with T scores in the clinical range on each scale. Abbreviations: std= standard deviation of the mean.

	mean	std	min	max	T >60	T >60%
Inattention	79.74	11.955	40	90	398	90.0
Hyperactivity/Impulsivity	72.87	16.388	40	90	315	71.3
Learning Problems	75.95	11.912	42	90	390	88.2
Executive Function	73.81	12.906	40	90	363	82.1
Aggression	62.95	17.268	34	90	205	46.4
Peer Relationships	71.94	17.973	44	90	290	65.6

The questionnaire also contains two validity scales that may indicate response bias, i.e. the rater tries to convey an overly positive or negative impression to secure a certain outcome (Conners, 2013). The validity scales indicated a possibly overly negative response style for 80 responses. Highly negative scores may indicate extreme problems in the rating domains or a negative bias of the rater, which may overestimate the child's difficulties. Analyses were carried out including and excluding ratings with high Negative Impression scores.

The Behavioral Rating Inventory of Executive Function (BRIEF) is a questionnaire about behaviours associated with executive function problems for parents of children aged 5 to 18 years (Gioia, Isquith, & Kenworthy, 2000). There are eight subscales measuring behaviour problems related to inhibition, shifting, emotional control, initiation, working memory, planning/organising, organisation of materials, and monitoring.

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The Strengths and Difficulties Questionnaire (SDQ) is a parent-rated scale for children aged 8 to 16 years. It provides ratings for emotional symptoms and prosocial behaviour as well as scores for problems related to behavioural conduct, hyperactivity/inattention, and peer relationships.

Community Detection

Community detection is an optimisation clustering method. Networks in the current analysis represented the child-by-child correlations across the 6 scales of the Conners 3 questionnaire. The community algorithm starts with each network node, i.e. child, in a separate community and then iteratively parcellates the network into communities to increase the quality index (Q) until a maximum is reached. The current study used the algorithm described by Rubinov and Sporns (Rubinov & Sporns, 2011) as implemented in the Brain Connectivity Toolbox (https://sites.google.com/site/bctnet/) version of August 2016. This algorithm is not deterministic and may yield different solutions at each run. In order to reach a stable community assignment, we applied the consensus clustering method described by Lancichinetti and Fortunato (Lancichinetti & Fortunato, 2012). In short, an average community assignment over 100 iterations was generated. The community assignment was then repeated further until the community assignment did not change between successive iterations. The analysis was implemented in Python 2.7.11. The code for the entire analysis is available online (http://www.github.com/joebathelt/).

Statistical Analysis

Groups defined by the community detection algorithm were compared on scales of the Conners 3 questionnaire. Shapiro-Wilk tests indicated that scores within groups deviated from normality assumptions (Shapiro & Wilk, 1965). Group contrasts were therefore based on non-parametric Mann-Whitney *U* tests (Mann & Whitney, 1947). The Bonferroni method was used to account for multiple comparisons. Statistical tests were carried out using Scientific Python (SciPy) version 0.17.0 implementation (Jones, Oliphant, Peterson, & others, 2001).

Structural connectome

The aim of this analysis was to explore whether the data-driven grouping was related to differences in brain structure. To this end, white matter connectivity of brain regions was estimated from diffusion-weighted images. Next, we employed a multivariate, dimension-reduction technique to relate the white-matter connectivity of brain regions to the group assignment.

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Participant sample for the connectome analyses

A subset of 191 families agreed to the neuroimaging part of the study. A total of 43 scans were excluded for poor quality, i.e. incomplete scan data, visually identified movement artefact, maximum displacement in the diffusion sequence above 3mm as determined by FSL eddy (see Figure 1 for an overview of attrition). The final sample consisted of 148 complete datasets (behaviour, T1, dwi). The MRI sample did not significantly differ in age from the behavioural sample (MRI sample [months]: mean=117.05, std=27.436, t(359)=1.34, p=0.181). The ratio of groups defined in the analysis of the behavioural sample was similar in the MRI subsample (MRI sample: C1: 0.36, C2: 0.33, C3: 0.30).

MRI data acquisition

Magnetic resonance imaging data were acquired at the MRC Cognition and Brain Sciences Unit, Cambridge U.K. All scans were obtained on the Siemens 3 T Tim Trio system (Siemens Healthcare, Erlangen, Germany), using a 32-channel quadrature head coil. The imaging protocol consisted of two sequences: T1-weighted MRI and a diffusion-weighted sequence.

T1-weighted volume scans were acquired using a whole brain coverage 3D Magnetisation Prepared Rapid Acquisition Gradient Echo (MP RAGE) sequence acquired using 1mm isometric image resolution. Echo time was 2.98 ms, and repetition time was 2250 ms.

Diffusion scans were acquired using echo-planar diffusion-weighted images with an isotropic set of 60 non-collinear directions, using a weighting factor of $b=1000s*mm^{-2}$, interleaved with a T2-weighted (b = 0) volume. Whole brain coverage was obtained with 60 contiguous axial slices and isometric image resolution of 2mm. Echo time was 90 ms and repetition time was 8400 ms.

Structural connectome construction

The white-matter connectome reconstruction followed the general procedure of estimating the most probable white matter connections for each individual and then obtaining measures of fractional anisotropy (FA) between regions (see Figure 2). The details of the procedure are described in the following paragraphs. In the current study, MRI scans were converted from the native DICOM to compressed NIfTI-1 format using the dcm2nii tool (http://www.mccauslandcenter.sc.edu/mricro/mricron/dcm2nii.html). Subsequently, a brain mask was derived from the b0-weighted volume of the diffusion-weighted sequence and the entire sequence was submitted for correction for participant movement and eddy current distortions through FSL's eddy tool. Next, non-local means de-noising (Coupe 2008) was applied using the Diffusion Imaging in Python (DiPy) v0.11 package (Garyfallidis 2014) to boost signal to noise ratio. The diffusion tensor model was fitted to the pre-processed images to derive maps of fractional anisotropy (FA) using dtifit from the FMRIB Software Library (FSL) v.5.0.6 (Behrens 2003). A spherical constrained deconvolution (CSD) model (Tournier 2008) was fitted to the 60-gradient-direction diffusion-weighted images using a maximum harmonic order of 8 using DiPy. An alternative analysis with a constant solid angle (CSA) model is present in the Supplementary Materials section. Next, probabilistic whole-brain tractography was

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

performed based on the CSD model with 8 seeds in any voxel with a General FA value higher than 0.1. The step size was set to 0.5 and the maximum number of crossing fibres per voxel to 2.

For ROI definition, T1-weighted images were preprocessed by adjusting the field of view using FSL's robustfov, non-local means denoising in DiPy, deriving a robust brain mask using the brain extraction algorithm of the Advanced Normalization Tools (ANTs) v1.9 (Avants 2011), and submitting the images to recon-all pipeline in FreeSurfer v5.3 (http://surfer.nmr.mgh.harvard.edu). Regions of interests (ROIs) were based on the Desikan-Killiany parcellation of the MNI template (Desikan 2006) with 34 cortical ROIs per hemisphere and 17 subcortical ROIs (brain stem, and bilateral cerebellum, thalamus, caudate, putamen, pallidum, hippocampus, amygdala, nucleus accumbens). The surface parcellation of the cortex was transformed to a volume using the aparc2aseg tool in FreeSurfer. Further, the cortical parcellation was expanded by 2mm into the subcortical white matter using in-house software. In order to move the parcellation into diffusion space, a transformation based on the T1-weighted volume and the b0-weighted image of the diffusion sequence was calculated using FreeSurfer's bbregister and applied to volume parcellation.

For each pairwise combination of ROIs, the number of streamlines intersecting both ROIs was estimated and transformed to a density map. A symmetric intersection was used, i.e. streamlines starting and ending in each ROI were averaged.

The weight of the connection matrices was based on fractional anisotropy (FA). To obtain FA-weighted matrices, the streamline density maps were binarized after thresholding and multiplied with the FA map and averaged over voxels to obtain the FA value corresponding to the connection between the ROIs. This procedure was implemented in-house based on DiPy v0.11 functions (Garyfallidis 2014). False positive streamline can introduce spurious results in structural connectome analyses (Zalesky et al., 2016). To remove spurious connections in the FA-weighted and streamline density-weighted networks, consensus thresholding was applied so that only connections that were present in more than 60% of the sample were retained (de Reus & van den Heuvel, 2013).

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

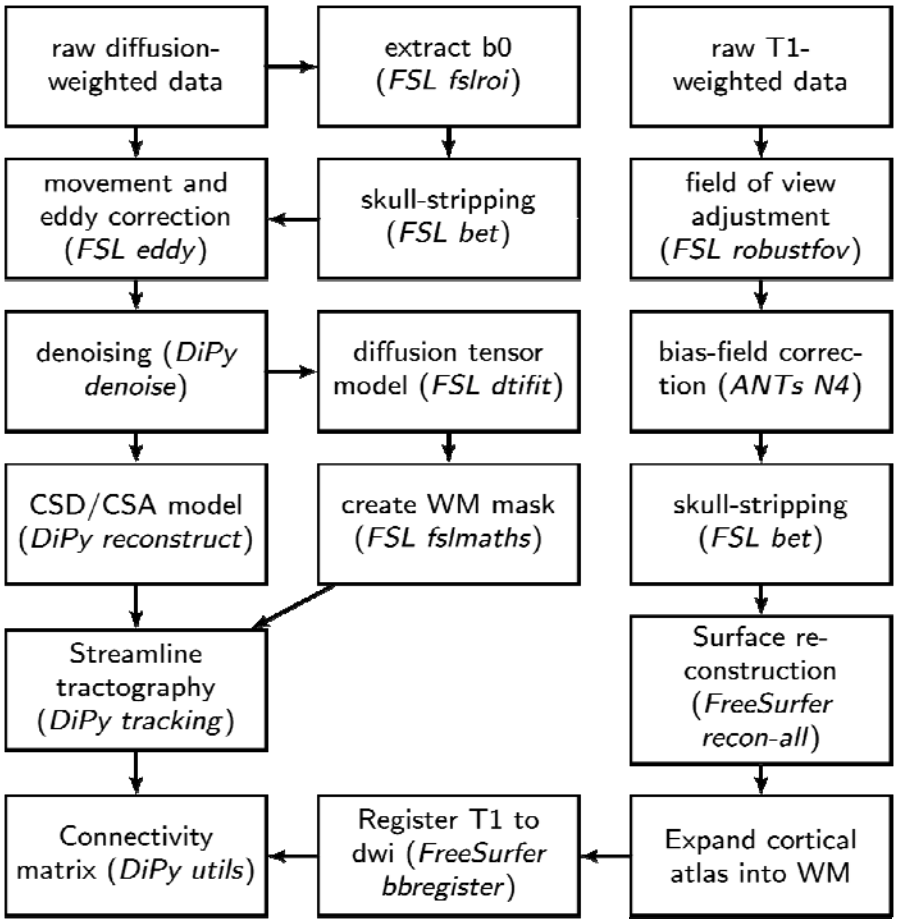


Figure 2: Overview of processing steps for structural connectome estimation

Statistical analysis of connectome data

For the analysis of the connectome data, the node degree of each node in the network was calculated for each participant. Partial least squares regression was used to identify the linear combination of brain areas that best explained group membership for the groups identified through community clustering. The PLS model was evaluated by fitting the model to random selection of 60% of the data and evaluating the model fit in a test set of 40%. The root-mean-squared error of a model based on the training data was significantly lower when assessed with the test data compared to randomly shuffled samples (10-fold cross-validated RMSE: mean=0.35, SE=0.025; permuted sample: mean=0.81, SE=0.018; permutation test: p=0.002).

The contribution of brain regions to the PLS latent variables was evaluated in a bootstrap procedure in which 60% of the sample was randomly selected and the PLS model was fitted (1000 permutations). The loading of brain regions onto PLS latent variables was expressed as the mean loading divided by the standard error across permutations (Krishnan, Williams, McIntosh, & Abdi, 2011). A Procrustes rotation was applied to align the factor across iterations of the permutation procedure. All procedures were

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

implemented using scikit-learn functions v0.18.1 under Python v2.7.12 (Pedregosa et al., 2011).

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Results

Community Detection indicates three subgroups

The current study employed graph theory to derive clusters of children with a similar profile across ratings on the Conners-3 questionnaire. The community detection algorithm in conjunction with consensus clustering arrived at a stable solution with three clusters. The quality index (Q=0.55) indicated strong separation of the clusters. A highly similar three cluster structure was also detected when excluding participants with a high negative impression rating (Q=0.59), and when randomly selecting half (Q=0.6) or a quarter of the sample (Q=0.61).

The cluster assignment resulted in roughly equal splits between the three clusters (Cluster 1: 150(33.93%), Cluster 2: 145(32.80%), Cluster 3: 147(33.25%); numbers for the whole sample). Statistical comparison between the groups indicated significant differences on all rating scales of the Conners 3 questionnaire between the groups (see Figure 3 and Table 3). Children in the clusters were characterised by problems associated with cognitive control (C1: Inattention, Hyperactivity/Impulsivity, Executive Function), learning difficulties (C2: Learning Problems), problems associated with deficits in behavioural conduct (C3: Aggression, Peer Relations). Standardised scores indicated that the majority of children in the current sample scored in the elevated to highly elevated range across all scales of the questionnaire compared to the normative sample of the Conners-3 questionnaire. The profiles based on scaled raw scores were also apparent when using the age-standardised scores. (see Figure 3b).

Next, the prevalence of pre-existing diagnoses in each cluster was evaluated. Children with a diagnosis of ADHD were over-represented in Cluster 1 (Executive Function) (see Table 4 for a breakdown of diagnoses per cluster, $X^2(3,354)=72.87$, $p=0.000$). Other diagnoses were equally distributed between the clusters (ASD: $X^2(3,354)=0.06$, $p=0.971$, Anxiety/Depression: $X^2(3,354)=0.54$, $p=0.764$, Learning Deficit: $X^2(3,354)=3.88$, $p=0.144$).

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

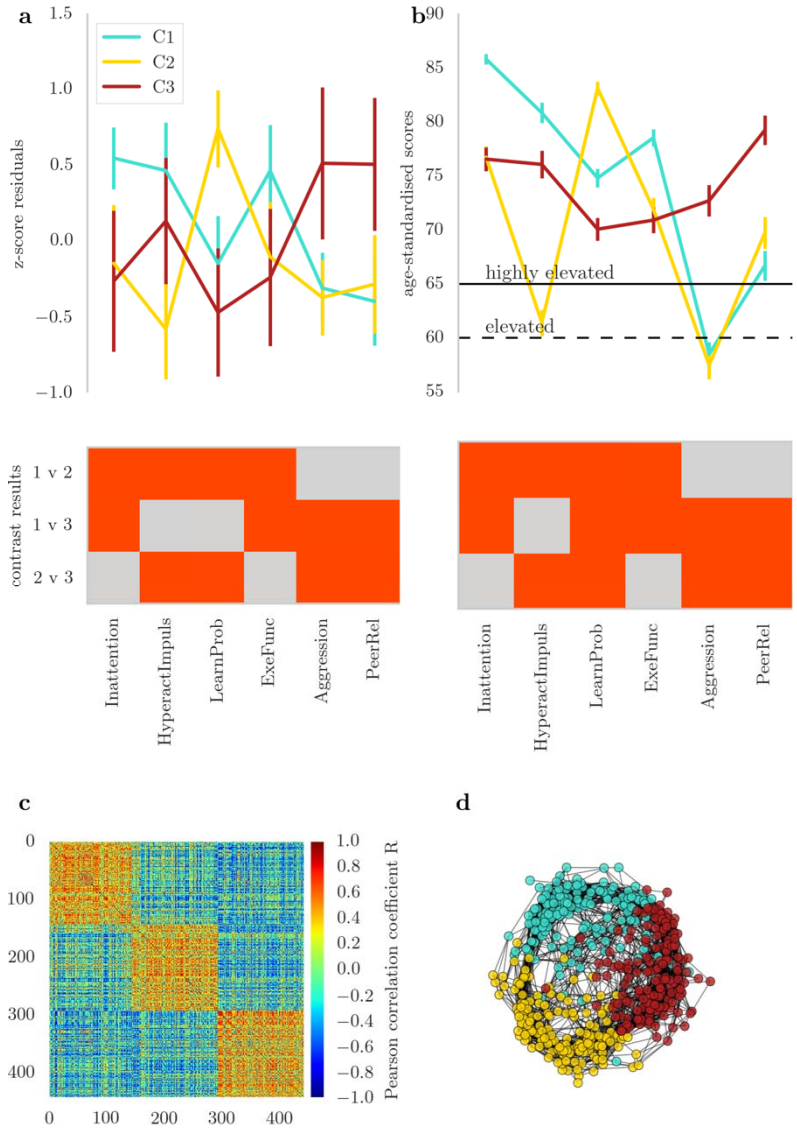


Figure 3: ***a*** *Profile of ratings on the Conners 3 questionnaire in the three clusters indicated by the community detection algorithm. The top of the figure shows the mean of scores in each group with two standard errors. The scores represent residuals after regressing the effect of age. The bottom figure shows the results of group-wise contrasts on each scale. Red indicates a significant difference between groups ( ) after Bonferroni correction.* ***b*** *Comparison of the groups on scores standardised with reference to the normative data of the Conners-3 questionnaire.* ***c*** *Child-by-child correlation matrix of Conners-3 scores after ordering the matrix according to the cluster assignment indicated by consensus clustering. The order matrix shows a clear separation between the clusters.* ***d*** *Correlation matrix in a spring layout colour-coded according to the cluster assignment indicated by consensus clustering. The spring layout representation shows clear spatial separation between the clusters.*

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 3: Scales of the Conners-3 questionnaire: Inattention, Hyperactivity/Impulsivity (HyperactImpuls.), Learning Problems (LearnProb.), Executive Function (ExeFunc), Aggression, Peer Relationship Problems (PeerRel.); mad: median absolute deviance; U: Mann Whitney U statistic; all p-values are Bonferroni corrected.

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

	Learning (1)		Exec. Function (2)		Conduct (3)		1 v 2		1 v 3		2 v 3	
	median	mad	median	mad	median	mad	U	*p*	U	*p*	U	*p*
Inattention	0.71	0.404	0.11	0.810	0.01	0.959	4867	p<0.001	6487	p<0.001	12026	1.000
HyperactImpuls.	0.63	0.636	-0.76	0.682	0.38	0.886	3149	p<0.001	9294	0.133	7321	p<0.001
LearnProb.	-0.15	0.642	0.90	0.524	-0.58	0.876	3269	p<0.001	9051	0.052	4137	p<0.001
ExeFunc	0.60	0.599	-0.10	0.760	-0.13	0.953	5427	p<0.001	7027	p<0.001	11782	1.000
Aggression	-0.53	0.452	-0.59	0.497	0.26	1.048	7649	0.843	6885	p<0.001	6871	p<0.001
PeerRel.	-0.56	0.571	-0.45	0.665	0.67	0.938	8085	1.000	5744	p<0.001	7099	p<0.001

Table 1: Scales of the Conners-3 questionnaire: Inattention, Hyperactivity/Impulsivity (HyperactImpuls.), Learning Problems (LearnProb.), Executive Function (ExeFunc), Aggression , Peer Relationship Problems (PeerRel.); mad: median absolute deviance; U: Mann Whitney U statistic; all p-values are Bonferroni corrected

Table 4: Breakdown of diagnoses in each cluster identified through data-driven clustering.

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

	C1	C2	C3	Total
ADHD	33	4	24	61
ASD	13	5	6	24
Learning Deficit	3	22	7	32
Other	7	8	8	23
None	94	106	102	302
Total	150	145	147	442

Subgroups show differences in other questionnaire measures of executive function and everyday difficulties

Next, the groups defined through community assignment based on Conners-3 data were compared on other questionnaire measures of behavioural problems linked to executive function difficulties (BRIEF) and everyday behavioural problems (SDQ). A comparison of these measures indicated significant differences between the groups. For the BRIEF, children in Cluster 1 (hyperactivity/executive problems) had more problems with working memory. Children in Cluster 2 (learning problems) were rated as having fewer difficulties with inhibition and monitoring and Cluster 3 (conduct problems) were also rated as having significantly higher problems in emotional control compared to the other groups (see Figure 4a).

For the strengths and difficulties questionnaire, children in Cluster 1 (hyperactivity/executive problems) were characterised by high ratings for hyperactivity compared to Cluster 2 (learning problems), but lower conduct and peer relationship problem ratings compared to Cluster 3 (conduct problems). Children in Cluster 2 (learning problems) received significantly lower ratings for problems related to hyperactivity. Children in Cluster 3 (conduct problems) received significantly higher ratings for conduct and peer relationship problems (see Figure 4b).

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

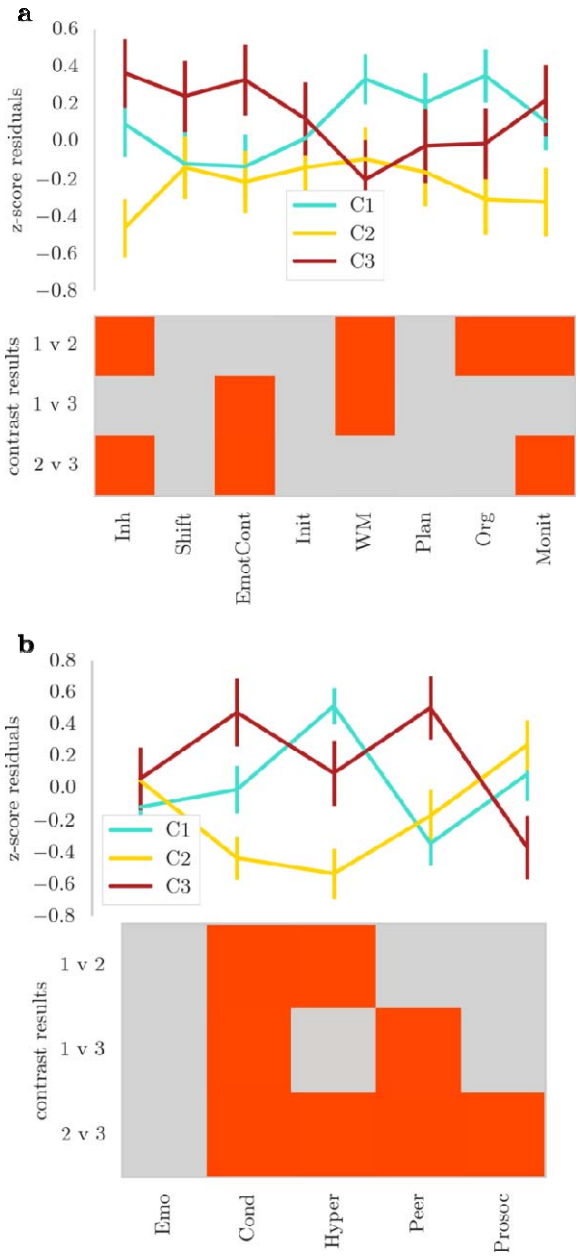


*Figure 4: Profile of ratings for children in the clusters defined by consensus module assignment on **a**) a questionnaire on executive function difficulties (BRIEF) and **b**) a questionnaire on strengths and difficulties (SDQ). The lines indicate the mean of each group across the questionnaire scales with error bars showing two standard errors around the mean. The bottom of each figure shows the binary outcome of t-tests comparing the groups. Red indicates a significant result ( ) after Bonferroni correction. Please note that higher scores indicate a higher level of difficulties on each scale, apart from the Prosocial Behaviour (Prosoc) scale where high scores indicate more prosocial behaviour. Abbreviations: top: Inh=Inhibition, EmotCont=Emotional Control, Init= Initiate, WM=Wokring Memory, Org=Organisation of Materials, Monit=Monitoring;*

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

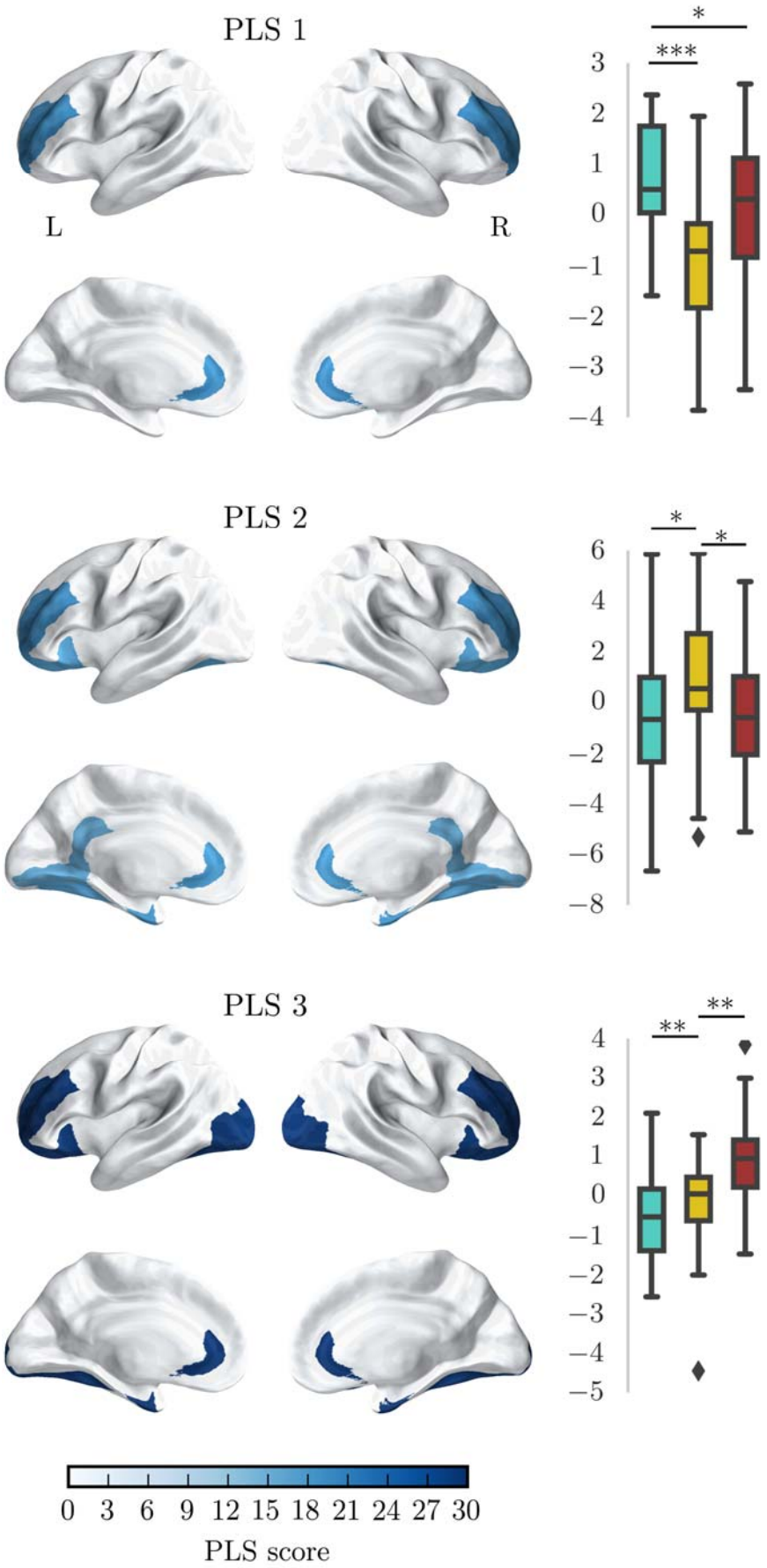
bottom: Emo=Emotional Problems, Cond=Conduct Problems, Hyper=Hyperactivity, Peer=Peer Problems, Prosoc=Prosocial Behaviour.

Subgroups show differences in the structural connectome

Next, we investigated the relationship between white matter connectivity and the groups defined through consensus clustering using partial least squares (PLS) regression. The first three PLS components explained 48% of variance in group membership (Component 1: 21.23% (SD: 4.302); Component 2: 16.28% (SD: 5.944); Component 3: 10.57% (SD: 4.277), bootstrapped mean and standard deviation (SD) over 1000 permutations). Further components explained less than 5% of variance and were therefore dropped from the analysis. Comparison of component loadings per group indicated significant lower loading of C1 (Inattention/Hyperactivity) compared to the other groups for PLS component 1, significantly higher loading in C1 (Inattention/Hyperactivity) compared to C3 (Aggression/Peer Problems) for PLS component 2, and significantly lower loading in C1 (Inattention/ Hyperactivity) compared to C2 (Learning Problems) for PLS component 3.

There were differences in the brain areas that distinguished the groups. PLS 1 that distinguished between C1 (Inattention/Hyperactivity) and the other groups loaded most heavily on the rostral middle frontal, superior frontal, lateral orbitofrontal, anterior cingulate, lateral occipital and fusiform cortex (see Figure 5). The second PLS component, which distinguished between C2 (Learning Problems) and C1 (Inattention/Hyperactivity), loaded the most on the rostral middle frontal, lateral orbitofrontal, anterior and posterior cingulate, and lateral occipital cortex. The third PLS component, which distinguished C3 (Conduct) from the other groups, loaded on the lateral orbitofrontal, anterior cingulate, and entorhinal cortex, and also on connections of the right pallidum and putamen (see Table 5).

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 5: Relationship between the node degree of brain regions in the structural connectome and clusters based on Conners-3 responses. The brain maps show the score of PLS components for brain regions that most strongly distinguished the group (top 20%). PLS scores above 2 are considered to be significantly predictive. The graphs show the statistical comparison of groups on loadings for each component.

Table 5: PLS scores for subcortical areas

		PLS 1	PLS 2	PLS 3
left	accumbens	0	0	0
	amygdala	0	0	0
	caudate	0	0	0
	hippocampus	0	0	0
	pallidum	0	0	0
	putamen	0	0	0
	thalamus	0	0	0
right	accumbens	0	0	0
	amygdala	0	0	0
	caudate	0	0	0
	hippocampus	0	0	0
	pallidum	0	0	35
	putamen	0	0	37
	thalamus	0	0	0

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Discussion

In the present study, we use a data-driven clustering algorithm to demonstrate that amongst a large sample of children referred by clinical and educational specialists there exist distinct behavioural profiles that transcend traditional diagnostic categories. Our data-driven clustering method produced groupings that were consistent in randomly-selected subsets of the sample and reliably reproduced in simulated data with a known structure even when adding considerable noise. Our method also identified behavioural profiles evident in other parent ratings that were not used to inform the original algorithm.

Specifically, we identified three subtypes of behavioural difficulties in children. One subgroup was characterised by high ratings for problems with behaviours related to inattention, hyperactivity/impulsivity, and executive function. This group was also rated as having increased difficulties with behaviours relating to working memory, organisation, planning, and hyperactivity on two other rating scales. The profile of behavioural difficulties in this subgroup captures core behavioural problems associated with the traditional ADHD diagnostic label (Barkley, 1997; Castellanos & Tannock, 2002; Pennington & Ozonoff, 1996, Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005). Indeed, a disproportion number of children with an ADHD diagnosis was assigned to this cluster. However, this subtype was not synonymous with ADHD as half of the children with an ADHD diagnoses were split across the other two clusters that displayed markedly different behavioural profiles.

Another subgroup identified through data-driven clustering were characterised by severe learning deficits relative to the other groups. In terms of their profiles on other questionnaires, this group received significantly lower ratings for behaviours relating to problems with inhibition. Scales relating to other aspects of attention and executive function were also lower relative to the other two groups. However, when comparing the groups to the standardisation sample of the questionnaire, children in the learning difficulties groups still scored in the elevated to clinical range. This indicates that children in this group fall below age expectation for attention and executive function, but have less pronounced difficulties than children with primary concerns in these areas. The profile of more problems with inattention and executive function combined with fewer problems with hyperactivity/impulsivity resembles the profile described for the inattentive subtype of ADHD (Carlson & Mann, 2000). However, the current results suggest that learning difficulties are a more distinguishing feature on parent questionnaires for these children, while concerns around aspects of executive function are not as pronounced.

A third subgroup was characterised by difficulties with aggression and peer relationships. Children in this group were also rated as having increased problems with behaviours related to emotional control and conduct on the two rating scales not used as part of the clustering algorithm. The distinction between groups with problems relating to either

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

executive function or behavioural conduct is reminiscent of the debate surrounding the overlap between ADHD and oppositional defiant disorder (ODD)/conduct disorder (CD). Some authors have argued for a high degree of overlap between these diagnostic groups (Szatmari, Boyle, & Offord, 1989), but evidence from genetic and imaging studies had suggested distinct pathophysiological mechanisms (Nadder, Rutter, Silberg, Maes, & Eaves, 2002, Rubia et al., 2009; Banaschewski et al., 2003). Consistent with these results, the current study shows that behavioural ratings of inattention/hyperactivity and aggression/peer relationship problems form distinct clusters.

These results demonstrate that data-driven clustering using a community detection algorithm can be used to characterise common and complex behavioural problems in children. The clustering algorithm identified groups that mirror some of the distinction of traditional diagnostic groups. One of the major advantages of this approach is that more homogeneous groupings were identified, and these are better suited for investigations into underlying biological mechanisms. Our data-driven sub-grouping was strongly associated with underlying differences in structural connectivity between groups. The areas that distinguish our groups have been suggested to play a role in relevant behaviours, making it possible to formulate hypotheses about neurobiological mechanisms associated with the different behavioural profiles. For instance, the group characterised by problems relating to attention and executive function showed differences in connectivity of the prefrontal, anterior cingulate cortex, and lateral occipital cortex. These differences in white matter connections of circuits related to inhibitory control (Miller & Cohen, 2001), goal-directed behaviour (Fjell et al., 2012) and visual attention (Castellanos & Proal, 2012) may play a role in the aetiology of these behavioural problems. In contrast, children with a profile of problems relating to emotional regulation and peer relationships were distinguished from the other groups by differences in white matter connectivity of the rostrolateral prefrontal cortex, anterior cingulate cortex, pallidum, and putamen. These findings may imply a difference in integration between the prefrontal cortex and the basal ganglia system (Finger et al., 2011; Rubia, 2011).

In summary, groups of children identified through data-driven clustering of executive dysfunction behaviours show different profiles of impairment that relate to conduct, learning, and executive difficulties, respectively. These groups were also distinguished by the connectivity of circuits previously implicated in executive function and behavioural regulation, including the prefrontal cortex, cingulate cortex, and their subcortical connections. These findings act as an important proof of principle: data-driven profiling provides an alternative means of distinguishing common and complex behavioural problems in children. This data-driven classification of behavioural problems may provide a better account of cognitive differences and relate more closely to neurobiological mechanisms than traditional diagnostic approaches.

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Acknowledgements

The Centre for Attention Learning and Memory (CALM) research clinic at the MRC Cognition and Brain Sciences Unit in Cambridge (CBSU) is supported by funding from the Medical Research Council of Great Britain to Duncan Astle, Susan Gathercole and Tom Manly. The clinic is led by Joni Holmes and managed by Francesca Woolgar. Data collection is assisted by a team of PhD students and researchers at the CBSU that includes Agnieszka Jaroslawska, Amy Johnson, Erica Bottacin, Erin Hawkins, Gemma Crickmore, Laura Forde, Sinead O'Brien, and Sara Gharooni. The authors wish to thank the many professionals working in children's services in the South-East and East of England for their support, and to the children and their families for giving up their time to visit the clinic.

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supplementary Analyses

Robustness of the consensus clustering algorithm

In order to test the reliability of the community detection algorithm under varying conditions, random networks with known community structure were created. The networks consisted of 100 nodes with 4 modules. The connection likelihood within and between clusters was systematically varied between 0.1 and 0.9. The quality index of the community structure was calculated at each combination of between- and within-cluster connection likelihood. The results indicated a high-quality index for network with higher within-cluster than outside-cluster connection likelihood (see Figure S1a). High connection density outside of clusters had a large influence, even when the connection likelihood within modules was very high.

For comparison with the empirical network of Conners-3 score correlations, the connection density within and between networks was calculated. To this end, all connections were binarized so that any connection with a Pearson correlation coefficient above 0 was set to 1. The connection density was estimated as the ratio between existing connections in the binarized empirical network and a fully connected network of the same size. Connection density within modules based on consensus clustering was 0.79 and connection density between modules was 0.05. Together with the results of the simulated networks, these connection densities indicate very high separation of the network clusters.

We further tested the robustness of the community assignment by adding increasing percentages of random Gaussian noise (=0, =1) to the network matrix and repeated the consensus clustering procedure (see Figure S1b). The quality index indicated good separation of the clusters between 5 and 30% noise (Q between 0.62 and 0.65). No stable assignment could be reached at 35% of noise and above. These results indicate that the community assignment is robust to a considerable amount of noise.

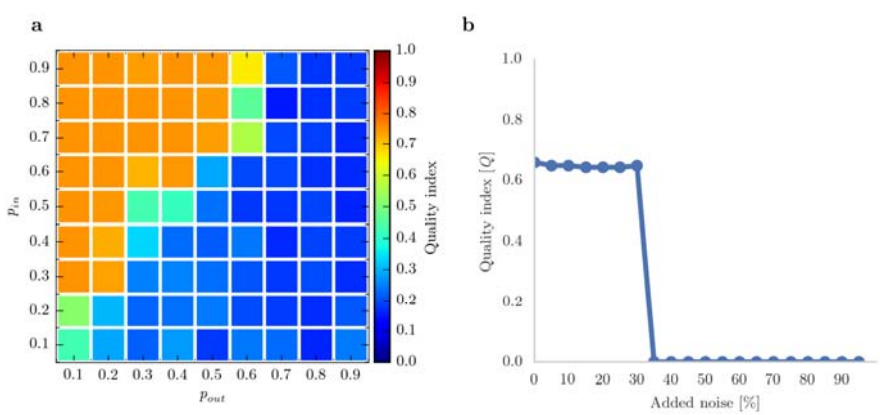


Figure S1: Results of robustness testing ***a****: Quality indices of consensus clustering using simulated networks with varying levels of within ( ) and between ( ) connections*

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

probabilities. High within-cluster and low between-cluster connectivity lead to high separation of clusters with consensus clustering, i.e. high quality-indices. ***b****: Consensus clustering using the empirical child-by-child network of Conners-3 correlations with varying levels of added noise. The three-cluster solution could be reconstructed up to 30% of added Gaussian noise. At higher-level of noise, no clustering solution could be obtained.*

Influence of connection weight and thresholding on structural connectome results

Different methods exist for the construction of structural networks from diffusion-weighted data and there is currently no scientific consensus on the best approach (Qi, Meesters, Nicolay, ter Haar Romeny, & Ossenblok, 2015). Networks in the current analysis were waited by fractional anisotropy (FA), a commonly used measure of white matter organisation based on the diffusion tensor model. FA characterises the directedness of diffusion within a voxel, but may lead to misinterpretation in regions of crossing fibres (Douaud et al., 2011). Therefore, the main analysis was repeated with networks weighted by Generalized FA (GFA) based on a constant solid angle (CSA) model, which is better able to take crossing fibres into account (Tuch, 2004). The node degree for each brain region was identical for the GFA and FA model for density thresholds between 5% and 15% (Kolmogorov-Smirnov test for two samples: p=1.0 uncorrected for all regions). It follows that the PLS analysis provides the same results for networks weighted by FA and GFA as this analysis was based on node degrees and node degrees were identical for both models within the relevant density range.

Another potential source of variation in the analysis is the density threshold. Network analyses are sensitive to the number of connections. Therefore, density thresholding is often applied, but the chosen threshold may influence the results of the analysis. For the current investigation, the influence of different density thresholds was systematically investigated by repeating the analysis over a range of densities and comparing the factor scores in a repeated-measures analysis of variance model with factors for density and the interaction between density and component (components loading density + component + density*component). The results indicated no significant effect of density or the interaction between density and any component (model fit: $F(9, 14790)<0.001$, $p=1$, Adjusted-R^2=-0.001; Density: $t<0.001$, $p=1$, Interactions: $t<0.001$, $p=1$).

References

Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A., & Gee, J. C. (2011). An open source multivariate framework for n-tissue segmentation with evaluation on public data. *Neuroinformatics*, *9*(4), 381–400. http://doi.org/10.1007/s12021-011-9109-y

Barabasi, A.-L. (2016). Communities. In *Network Science* (pp. 321–377). Cambridge, UK: Cambridge University Press.

Barkley, R. A. (1997). Behavioral inhibition sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65–94. http://doi.org/10.1037/0033-2909.121.1.65

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Banaschewski, T., Brandeis, D., Heinrich, H., Albrecht, B., Brunner, E., & Rothenberger, A. (2003). Association of ADHD and conduct disorder - brain electrical evidence for the existence of a distinct subtype. *Journal of Child Psychology and Psychiatry*, *44*(3), 356–376. http://doi.org/10.1111/1469-7610.00127

Bathelt, J., Gathercole, S. E., Johnson, A., Astle, D. E. (in press). Differences in brain morphology and working memory capacity across childhood. Developmental Science. http://doi.org/10.1111/desc.12579

Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., et al. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, *50*(5), 1077–1088. http://doi.org/10.1002/mrm.10609

Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186–198. http://doi.org/10.1038/nrn2575

Carlson, C. L., & Mann, M. (2000). Attention-deficit/hyperactivity disorder, predominately inattentive subtype. *Child and Adolescent Psychiatric Clinics of North America*.

Castellanos, F. X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. *Nature Reviews Neuroscience*, *3*(8), 617–628. http://doi.org/10.1038/nrn896

Castellanos, F. X., & Proal, E. (2012). Large-scale brain systems in ADHD: beyond the prefrontal-striatal model. *Trends in Cognitive Sciences*, *16*(1), 17–26. http://doi.org/10.1016/j.tics.2011.11.007

Clayden, J. D., Jentschke, S., Munoz, M., Cooper, J. M., Chadwick, M. J., Banks, T., & Vargha-Khadem, F. (2011). Normative development of white matter tracts: similarities and differences in relation to age gender, and intelligence. *Cerebral Cortex*, *22*(8), 1738–1747. http://doi.org/10.1093/cercor/bhr243

Corbett, B. A., Constantine, L. J., Hendren, R., Rocke, D., & Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder, and typical development. *Psychiatry Research*, *166*(2-3), 210–222. http://doi.org/10.1016/j.psychres.2008.02.005

Collin, G., & van den Heuvel, M. P. (2013). The ontogeny of the human connectome: development and dynamic changes of brain connectivity across the life span. *The Neuroscientist: a Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *19*(6), 616–628. http://doi.org/10.1177/1073858413503712

Conners, K. C. (2013). *Conners 3rd edition Parent Short-Form*. North Tonawanda, NY: Multi-Health Systems Inc.

Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., & Barillot, C. (2008). An Optimized Blockwise Nonlocal Means Denoising Filter for 3-D Magnetic Resonance Images. *IEEE Transactions on Medical Imaging*, *27*(4), 425–441. http://doi.org/10.1109/tmi.2007.906087

Czamara, D., Tiesler, C. M. T., Kohlbock, G., Berdel, D., Hoffmann, B., Bauer, C.-P., et al. (2013). Children with ADHD Symptoms Have a Higher Risk for Reading, Spelling and Math Difficulties in the GINIplus and LISAplus Cohort Studies. *PloS One*, *8*(5), e63859. http://doi.org/10.1371/journal.pone.0063859

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. http://doi.org/10.1016/j.neuroimage.2006.01.021

de Reus, M. A., & van den Heuvel, M. P. (2013). Estimating false positives and negatives in brain networks. *NeuroImage*, *70*, 402–409. http://doi.org/10.1016/j.neuroimage.2012.12.066

Douaud, G., Jbabdi, S., Behrens, T. E. J., Menke, R. A., Gass, A., Monsch, A. U., & Smith, S. (2011). DTI measures in crossing-fibre areas: Increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimers disease. *NeuroImage*, *55*(3), 880–890. http://doi.org/10.1016/j.neuroimage.2010.12.008

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Finger, E. C., Marsh, A. A., Blair, K. S., Reid, M. E., Sims, C., Ng, P., & Blair, R. J. R. (2011). Disrupted reinforcement signaling in the orbitofrontal cortex and caudate in youths with conduct disorder or oppositional defiant disorder and a high level of psychopathic traits. *American Journal of Psychiatry*, *168*(2), 152–162. http://doi.org/10.1176/appi.ajp.2010.10010129

Fjell, A. M., Walhovd, K. B., Brown, T. T., Kuperman, J. M., Chung, Y., Hagler, D. J., & Dale, A. J. (2012). Multimodal imaging of the self-regulating developing brain. *Proceedings of the National Academy of Sciences*, *109*(48), 19620–19625. http://doi.org/10.1073/pnas.1208243109

Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., et al. (2014). Dipy, a library for the analysis of diffusion MRI data. *Frontiers in Neuroinformatics*, *8*(175), 8. http://doi.org/10.3389/fninf.2014.00008

Germano, E., Gagliano, A., & Curatolo, P. (2010). Comorbidity of ADHD and Dyslexia. *Developmental Neuropsychology*, *35*(5), 475–493. http://doi.org/10.1080/87565641.2010.494748

Gioia, G. A., Isquith, P. K., & Kenworthy, L. (2000). Behavior Rating Inventory of Executive Function. Lutz, FL: PAR Inc.

Gjone, H., Stevenson, J., & Sundet, J. M. (1996). Genetic influence on parent-reported attention-related problems in a Norwegian general population twin sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*(5), 588–96– discussion 596–8. http://doi.org/10.1097/00004583-199605000-00013

Harty, S. C., Miller, C. J., Newcorn, J. H., & Halperin, J. M. (2008). Adolescents with Childhood ADHD and Comorbid Disruptive Behavior Disorders: Aggression, Anger, and Hostility. *Child Psychiatry and Human Development*, *40*(1), 85–97. http://doi.org/10.1007/s10578-008-0110-0

Jones, E., Oliphant, T., Peterson, P., & others. (2001–2001–). SciPy: Open source scientific tools for Python.

Krishnan, A., Williams, L. J., McIntosh, A. R., & Abdi, H. (2011). Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. *NeuroImage*, *56*(2), 455–475. http://doi.org/10.1016/j.neuroimage.2010.07.034

Lancichinetti, A., & Fortunato, S. (2012). Consensus clustering in complex networks. *Sci. Rep.*, *2*. http://doi.org/10.1038/srep00336

Lebel, C., Walker, L., Leemans, A., Phillips, L., & Beaulieu, C. (2008). Microstructural maturation of the human brain from childhood to adulthood. *NeuroImage*, *40*(3), 1044–1055. http://doi.org/10.1016/j.neuroimage.2007.12.053

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.*, *18*(1), 50–60. http://doi.org/10.1214/aoms/1177730491

Marcus, D. K., Barry, T. D. (2011). Does Attention-Deficit/Hyperactivity Disorder Have a Dimensional Latent Structure? A Taxometric Analysis. *Journal of Abnormal Psychology*, *120*(2), 427–442. http://doi.org/10.1037/a0021405

Miller, E. K., & Cohen, J. D. (2001). An integrated theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. http://doi.org/10.1146/annurev.neuro.24.1.167

Mrug, S., Molina, B. S. G., Hoza, B., Gerdes, A. C., Hinshaw, S. P., Hechtman, L., & Arnold, L. E. (2012). Peer Rejection and Friendships in Children with Attention-Deficit/Hyperactivity Disorder: Contributions to Long-Term Outcomes. *Journal of Abnormal Child Psychology*, *40*(6), 1013–1026. http://doi.org/10.1007/s10802-012-9610-2

Nadder, T. S., Rutter, M., Silberg, J. L., Maes, H. H., & Eaves, L. J. (2002). Genetic effects on the variation and covariation of attention deficit-hyperactivity disorder (ADHD) and oppositional-defiant disorder/conduct

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

disorder (ODD/CD) symptomatologies across informant and occasion of measurement. *Psychological Medicine*, *32*(01). http://doi.org/10.1017/s0033291701004792

Nigg, J.T. (2006). What causes ADHD? Toward a multi- path model for understanding what goes wrong and why (p. 422). New York: The Guilford Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pennington, B. F., & Ozonoff, S. (1996). Executive Functions and Developmental Psychopathology. *Journal of Child Psychology and Psychiatry*, *37*(1), 51–87. http://doi.org/10.1111/j.1469-7610.1996.tb01380.x

Qi, S., Meesters, S., Nicolay, K., ter Haar Romeny, B. M., & Ossenblok, P. (2015). The influence of construction methodology on structural brain network measures: A review. *Journal of Neuroscience Methods*, *253*, 170–182. http://doi.org/10.1016/j.jneumeth.2015.06.016

Rubia, K., Smith, A. B., Halari, R., Matsukura, F., Mohammad, M., Taylor, E., & Brammer, M. J. (2009). Disorder-specific dissociation of orbitofrontal dysfunction in boys with pure conduct disorder during reward and ventrolateral prefrontal dysfunction in boys with pure ADHD during sustained attention. *American Journal of Psychiatry*, *166*(1), 83–94. http://doi.org/10.1176/appi.ajp.2008.08020212

Rubia, K. (2011). Cool inferior frontostriatal dysfunction in attention-deficit/hyperactivity disorder versus hot ventromedial orbitofrontal-limbic dysfunction in conduct disorder: A review. *Biological Psychiatry*, *69*(12), e69–e87. http://doi.org/10.1016/j.biopsych.2010.09.023

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, *52*(3), 1059–1069. http://doi.org/10.1016/j.neuroimage.2009.10.003

Rubinov, M., & Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. *NeuroImage*, *56*(4), 2068–2079. http://doi.org/10.1016/j.neuroimage.2011.03.069

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3-4), 591–611. http://doi.org/10.1093/biomet/52.3-4.591

Sonuga-Barke, E. J. S. (1998). Categorical Models of Childhood Disorder: A Conceptual and Empirical Analysis. *Journal of Child Psychology and Psychiatry*, *39*(1), 115–133. http://doi.org/10.1017/S0016774600023726

Sonuga-Barke, E. J. S., & Halperin, J. M. (2010). Developmental phenotypes and causal pathways in attention deficit/hyperactivity disorder: potential targets for early intervention? *Journal of Child Psychology and Psychiatry*, *51*(4), 368–389. http://doi.org/10.1111/j.1469-7610.2009.02195.x

Stevens, M. C., Skudlarski, P., Pearlson, G. D., & Calhoun, V. D. (2009). Age-related cognitive gains are mediated by the effects of white matter development on brain network integration. *NeuroImage*, *48*(4), 738–746. http://doi.org/10.1016/j.neuroimage.2009.06.065

Szatmari, P., Boyle, M., & Offord, D. R. (1989). ADDH and conduct disorder: Degree of diagnostic overlap and differences among correlates. *Journal of the American Academy of Child & Adolescent Psychiatry*, *28*(6), 865–872. http://doi.org/10.1097/00004583-198911000-00010

Tuch, D. S. (2004). Q-ball imaging. *Magnetic Resonance in Medicine*, *52*(6), 1358–1372. http://doi.org/10.1002/mrm.20279

Tournier, J. D., Yeh, C.-H., Calamante, F., Cho, K.-H., Connelly, A., & Lin, C.-P. (2008). Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data. *NeuroImage*, *42*(2), 617–625. http://doi.org/10.1016/j.neuroimage.2008.05.002

bioRxiv preprint doi: https://doi.org/10.1101/158949; this version posted July 5, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, *57*(11), 1336–1346. http://doi.org/10.1016/j.biopsych.2005.02.006

van den Heuvel, M. P., Kahn, R. S., Goni, J., & Sporns, O. (2012). High-cost, high-capacity backbone for global brain communication. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(28), 11372–11377. http://doi.org/10.1073/pnas.1203593109

World Health Organization (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*.

Zalesky, A., Fornito, A., Cocchi, L., Gollo, L. L., van den Heuvel, M. P., & Breakspear, M. (2016). Connectome sensitivity or specificity: which is more important? *NeuroImage*, *142*, 407–420. http://doi.org/10.1016/j.neuroimage.2016.06.035