**Supplementary Information**

**Brain-behavior patterns define a dimensional biotype in medication-naïve adults with attention-deficit hyperactivity disorder**

Hsiang-Yuan Lin, MD1,2+, Luca Cocchi, PhD2+, Andrew Zalesky, PhD3, Jinglei Lv, PhD2, Alistair Perry, PhD2, Wen-Yih Isaac Tseng, MD, PhD4,5, Prantik Kundu, PhD6, Michael Breakspear, MBBS, PhD, FRANZCP2,7, Susan Shur-Fen Gau, MD, PhD1,5\*

**1. Supplementary Methods**

**1.1. Data**

**1.1.1. Measures for ADHD symptoms**

*1.1.1.1. The Adult ADHD Self-Report Scales*

The Adult ADHD Self-Report Scales (ASRS), an 18-question scale, was developed in conjunction with the revision of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). The ASRS consists of two subscales, Inat­tention (nine items) and Hyperactivity-Impulsivity (nine items), according to the 18 *DSM-IV* ADHD symptom crite­ria. Each item asks how often a symptom occurred during the last 6 months on a 5-point Likert scale: 0=*never*, 1=*rarely*, 2=*sometimes*, 3=*often*, and 4=*very often*. The psychometric properties of the Chinese ASRS have been established in a sample of 4,329 Taiwanese young adults (Yeh *et al.*, 2008). The intraclass correlations (ICCs) for test-retest reliability ranged from 0.80 for the Inattention sub­scale, 0.82 for the Hyperactivity-Impulsivity subscale, and .85 for the total score. The internal consistency (Cronbach’s α) was high for the Inattention subscale (0.87), the Hyperac­tivity-Impulsivity subscale (0.85), and the total score (0.91). It has been used in studies on adult ADHD and sleep problems, anxiety/depression symptoms, and quality of life in Taiwan (Gau *et al.*, 2007; Chao *et al.*, 2008).

*1.1.1.2. The Swanson, Nolan, and Pelham, Version IV Scale (SNAP-IV)-Parent form*  
The SNAP-IV is a 26-item rating instrument including the core DSM-IV-derived ADHD subscales of IA, HI and OD subscales (items 1-9, 10-18, and 19-26, respectively) (Swanson *et al.*, 2001). Each item is rated on a 4-point Likert scale, 0-3 for “not at all”, “just a little”, “quite a lot”, and “very much” based on and parents’ report. The norm and psychometric properties of the Chinese version of SNAP-IV have been well established in Taiwan by Gau and colleagues (Gau *et al.*, 2008). The scale has good test-retest reliability (ICCs 0.59~0.72), high internal consistency (Cronbach’s α>0.88) and discriminative validity (Gau *et al.*, 2008) and is commonly used in clinical evaluation and research in Taiwanese child and adolescent populations (Yang *et al.*, 2013).

*1.1.1.3. The modified adult version of the ADHD supplement of the Chinese version of the Schedule for Affective Disorders and Schizophrenia–Epidemiological Version (K-SADS-E)*

The K-SADS-E is a semi-structured interview scale for the systematic assessment of both past and current episodes of mental disorders in children and adolescents (Orvaschel *et al.*, 1982). Development of the Chinese K-SADS-E was completed by the Child Psychiatry Research Group in Taiwan (Gau and Soong, 1999). This included a two-stage translation and modification for several items with psycholinguistic equivalents relevant to the Taiwanese culture and further modification to meet the DSM-IV diagnostic criteria, with high reliability (generalized kappa coefficients ranging from 0.73 to 0.96 for all mental disorders) and validity (sensitivity 78% and specificity 98%) (Gau *et al.*, 2005). In order to obtain the information about ADHD symptoms and diagnoses in adulthood according to the DSM-IV diagnostic criteria, semi-structured interviews were conducted using both the modified adult ADHD supplement and the Conners' Adult ADHD Diagnostic Interview for DSM-IV (Takahashi *et al.*, 2014). The results showed that the ADHD diagnosis in childhood and current adulthood based on the two clinical instruments achieved total agreement (i.e., people who had been diagnosed with ADHD in childhood and/or current adulthood using the modified adult ADHD supplement of the K-SADS-E also acquired the ADHD diagnosis based on the Conners' Adult ADHD Diagnostic Interview).

**1.2. Analyses**

**1.2.1. Multi-echo independent component analysis (ME-ICA)**

ME-ICA initially decomposed multi-echo rs-fMRI data into independent components using FastICA (Hyvarinen, 1999b). Independent components were subsequently categorized as BOLD or non-BOLD components based on Kappa and Rho values, which were yielded from signal models reflecting the BOLD-like or non-BOLD-like signal decay processes (Kundu *et al.*, 2012). BOLD-related signals show linear dependence of percent signal changes on TE, which is the characteristic of the T2\* decay (Huettel *et al.*, 2008). On the other hand, non-BOLD signal amplitudes demonstrate TE-independence. TE dependence of BOLD signal was measured using the pseudo-F-statistic Kappa, with components that scaled strongly with TE having high Kappa scores. Non-BOLD components were identified by TE independence measured by the pseudo-F-statistic Rho. By removing non-BOLD components, data were denoised for head motion, physiological, and scanner artifacts (Kundu *et al.*, 2013).

**1.2.2. Functional network connectivity analysis**

*1.2.2.1. Independent Component Analysis (ICA)*

After preprocessing, the temporally concatenated probabilistic ICA algorithm (temporally concatenated) implemented in FSL MELODIC (Beckmann and Smith, 2004) was used to analyze the rs-fMRI data of all participants. Non-brain voxels were masked with voxel-wise demeaning of the data and normalization of the voxel-wise variance. Next, the processed data were whitened and projected into a 20-dimensional subspace using a Principal Components Analysis (PCA). This step provided a fine-grained decomposition of interconnected brain regions (Smith *et al.*, 2009). These whitened observations were decomposed into sets of vectors that describe signal variation across (i) the temporal domain (time courses), (ii) the session/subject domain, and (iii) the spatial domain (spatial maps). This decomposition was implemented through a non-Gaussian spatial source distribution using a fixed-point iteration technique (Hyvarinen, 1999a). Estimated component maps were divided by the standard deviation of the residual noise, with a threshold of 0.5 set (the probability that needed to be exceeded by a voxel to be considered ‘active’ in the component of interest) by fitting a mixture model to the histogram of intensity values (Beckmann and Smith, 2004).

We selected resting-state networks according to their known spatial distribution (Smith *et al.*, 2009; Yeo *et al.*, 2011; Cocchi *et al.*, 2012). We extracted 20 ICA components, 14 of which are consistently identified as canonical resting-state networks (Yeo *et al.*, 2011). The similarity of these 14 resting-state networks with those previously identified was quantified using spatial correlation (all spatial correlation values >0.4) and confirmed by visual inspection. Only these 14 components (networks) were considered in subsequent analyses (Supplementary Fig. 2A).

*1.2.2.2. Functional Network Connectivity*

The summary time-course for each resting-state network was calculated at the individual participant’s level by spatial regression of the full set of 20 ICA components against each participant’s denoised rs-fMRI data. This approach models are overlapping variance to account for the potential effects of residual noise captured by the non-physiological valid components (N=6). We calculated functional network connectivity (FNC) (Jafri *et al.*, 2008; Lv *et al.*, 2016) using the Pearson correlation coefficient between each other summary time course. This resulted in a 3D FNC matrix with the dimensions of 14 × 14 (networks) × 203 (participants). Group differences in FNC were tested for each pair of networks using one-way analysis of variance (ANOVA), and FNC with significant group differences were further tested by 2-sample t-test to determine the direction of the difference. The significance threshold was set at *q*<0.05, corrected for multiple comparisons using false discovery rate (FDR) (Benjamini and Hochberg, 1995).

**1.2.3. Principal component analysis for ADHD symptoms**

To circumvent reporting biases (Asherson *et al.*, 2016), core ADHD symptoms were encapsulated as factor scores (DiStefano *et al.*, 2009) derived from a principal component analysis (PCA) of self-, parents-, and clinician-reported measures, including self-rated Adult ASRS (Yeh *et al.*, 2008), parent-rated SNAP-IV (Gau *et al.*, 2008), as well as a clinician-rated modified adult version of the ADHD supplement of the Chinese version of the K-SADS-E (Chang *et al.*, 2013; Ni *et al.*, 2013; Ni *et al.*, 2017) (the number of ADHD measures used in the PCA was 3). Two principal components were extracted, which explained 88.25% of the total variance. Factors were orthogonalized using Varimax rotation. Among them, the first component explained 51.58% of the total variance, and all of the hyperactivity-impulsivity subscales from the above three measures were consistently loaded on this component. The second component explained 36.67% of the total variance, and scores of inattention subdomain across 3 measures were loaded on the component. The principal component analysis was implemented using IBM SPSS Statistics for Macintosh, Version 22.0 (IBM Corp., Armonk, NY, USA).

|  |  |  |
| --- | --- | --- |
| *Symptoms scores patterns loaded onto two components (rotated component matrix)* | | |
|  | Component | |
| 1 (Hyperactivity-impulsivity) | 2 (Inattention) |
| SNAP\_Inattention | 0.244 | 0.743 |
| SNAP\_Hyperactivity-impulsivity | 0.842 | 0.182 |
| K-SADS-E\_Inattention | 0.224 | 0.806 |
| K-SADS-E\_ Hyperactivity-impulsivity | 0.798 | 0.172 |
| ASRS-A | 0.227 | 0.871 |
| ASRS-B | 0.754 | 0.416 |

**1.2.4. Canonical correlation analysis (CCA)**

Both connectivity and behavioral measures were normalized and demeaned. A further regression of in-scanner head motion confounds also performed following the approach of Smith and colleagues (Smith *et al.*, 2015) (<http://www.fmrib.ox.ac.uk/analysis/HCP-CCA>). To avoid overfitting the CCA, a PCA was undertaken using the FSLNets toolbox (Smith *et al.*, 2014) to reduce the dimensionality of the deconfounded functional connectivity matrix to three eigenvectors (explaining 31.83% of the total variance in the connectivity matrix; Supplementary Fig. 5). The data was reduced to this resolution to keep the methodological steps as per Smith et al. (Smith *et al.*, 2015), given the three behavioral measures selected in the CCA. We note that no consensus exists for component number selection (Abdi and Williams, 2010). Thus, we also employed a confirmatory CCA analysis based on a larger dimensionality of 5 eigenvectors (explaining 43.9% of the variance in the connectivity matrix). The primary (3 eigenvectors, *r*=0.430, FWE-corrected *p*=0.037) and confirmatory CCA (*r*=0.446, FWE-corrected *p*=0.049; Supplementary Table 4b) yielded similar results. Thus, only results from the primary CCA are reported in the main text.

We next assessed which functional connections were most strongly expressed by variations in the original sets of connections captured by each CCA mode. CCA provides an output vector describing the extent (weight) to which a given individual's connectivity pattern correlated with the CCA mode. We correlated this vector against the original connectivity matrix identified by the NBS analysis to obtain a vector mapping the relative weights and directional signs of the association between resting-state connectivity and the CCA mode (weighted feature vector). In line with what previously done, the strongest (top 25%) absolute values in this vector were retained to define the strongest associations between individual connectivity weights and behavioral measures (Smith *et al.*, 2015).

**1.2.5. Clustering algorithms for categorically subtyping ADHD**

To test the existence of ADHD categorical biotypes, we implemented several complementary analyses using the connectivity and clinical features derived from the significant CCA mode, and combined features from connectivity and clinical symptoms, respectively.

*1.2.5.1. k-means clustering algorithm based on brain-behavior features derived from the significant CCA mode.*

To assess whether the brain-behavior associations identified by the CCA could be clustered into non-overlapping subgroups, we first used *k*-means clustering on the features linearly projected by the CCA. This standard clustering procedure uses individual brain-behavior associations to assign each participant to exactly one of *k* clusters (based on clinical ADHD subtypes, a *k*=2 or 3 was used here) (Venkataraman *et al.*, 2009). To reach stable clustering results, for each setting of *k*, clustering was repeated for 10,000 times so that the participants-to-centroid distances within-cluster sum-of-squares were minimized.

*1.2.5.2. Multi-view spectral clustering algorithm based on features of functional connectivity and clinical symptoms*

With regards to multi-view spectral clustering algorithm (Shi and Malik, 2000; Kumar and Daumé, 2011), we considered clusters derived from the analysis of altered functional connectivity in ADHD compared to controls and features related to clinical symptoms/IQ as two views contributing to the clustering. Using the multi-view spectral clustering framework, the substantial variability of categorical subgrouping across multimodal features (connectivity and behavior) could be modeled and accounted for. This novel clustering method has the advantage of effectively addressing heterogeneity in the considered features by maximizing the agreement across multimodal clusters (Shi and Malik, 2000; Kumar and Daumé, 2011).

Spectral clustering uses connectivity (*denoised NBS results*) and clinical features (*inattention, hyperactivity-impulsivity, and IQ*), respectively, to generate two graphs. Nodes within the graphs represent individuals with ADHD whereas the edges represent the similarities between nodes (individuals). The two graphs (one mapping connectivity and one mapping behavior) were then partitioned using the normalized cut strategy, in which the top *k* eigenvectors of the normalized graph Laplacian, which carries the most discriminative information, are adopted to cut the graphs into clusters efficiently. Subsequent co-training algorithms search for target clusters that predict same labels for co-occurring patterns in each view. The spectral clustering algorithm of bi-partitioning sub-graph stopped when the normalized cut value (representing the similarity between the subjects within each possible cluster) is larger than the pre-set threshold. There is no consensus regarding the optimal threshold to be used. Thus, we examined thresholds ranging from 0.2 to 0.9 (incremental of 0.1) (Chen *et al.*, 2013). We used 10,000 iterations for co-training algorithms to converge on stable clusters (permuting for each threshold).

*1.2.5.3. Validity of k-means clustering*

We verify the validity of *k*-means clustering using average silhouette width values (Kononenko and Kukar, 2007), the Jaccard similarity (Hennig, 2008), and the gap statistic (Tibshirani *et al.*, 2001). This information is provided in Supplementary Table 6 (average silhouette width values and the Jaccard similarity) and Supplementary Fig. 7 (the Gap statistic).

The silhouette width value is a combination measure assessing intra-cluster homogeneity and inter-cluster separation. It is calculated by measuring how similar that point is to points in its own cluster when compared to points in other clusters. The cutoffs to interpret the validity of *k*-means clustering based on average silhouette width values are as follows (Kononenko and Kukar, 2007):

|  |
| --- |
| 0.71-1.0 A strong structure has been found. |
| 0.51-0.70 A reasonable structure has been found. |
| 0.26-0.50 The structure is weak and could be artificial. Try additional methods of data analysis. |
| <0.25 No substantial structure has been found. |

Jaccard’s similarity (Hennig, 2008) is defined as the size of the intersection divided by the size of the union of the assigned clusters and the resulting partitions from resampling pipelines. It allows estimating the frequency with which similar clusters were recovered in the data. The clustering results with Jaccard’s similarity <0.5 are considered unstable (Hennig, 2008).

The gap statistic (Tibshirani *et al.*, 2001) standardizes the graph of log(*Wk*), where *Wk* is the within-cluster dispersion defined by the within-cluster sum of squares around the cluster means, by comparing it to its expectation under an appropriate null reference distribution of the data. The ’*k*’ is the number of clusters. The estimate of the optimal number of clusters is defined by searching for the local maximum of the graph, and selecting the smallest *k* within one standard error of the local max.

*1.2.5.4. The issue of sample size for clustering analyses*

There is no clear indication regarding the minimum sample size necessary for clustering analyses. However, it is suggested that the minimal sample size for clustering analyses should not be less than 2m cases (*m*=number of features used), with 5\*2m considered preferable (Dolnicar, 2002). In the present study, we fed features linearly projected by the CCA (i.e., 1 for the brain connectivity feature; 1 for the behavior feature) into *k*-means clustering. That is, the minimum sample size for *k*-means clustering is 20 subjects (i.e., 5\*22=20). Concerning the multi-view spectral clustering, there has been very limited prior work investigating the minimum sample size required to obtain meaningful clusters. The multi-view spectral clustering algorithm is, however, considered robust for the high dimensionality and small-sample-size problem (Tao *et al.*, 2014). Indeed, a smaller sample size is generally required to obtain a solution (i.e., the most robust clustering results) using multi-view spectral clustering compared to single-view clustering (Kumar and Daumé, 2011).

In keeping with the above, the current sample size (N=80 ADHD) is appropriate for both *k*-means and spectral clustering (Dolnicar, 2002; Kumar and Daumé, 2011; Tao *et al.*, 2014).

**2. Supplementary Tables**

**Supplementary Table 1a.** Demographics among attention-deficit hyperactivity subtypes (ADHD) (based on the current presentation of ADHD psychopathology)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mean (SD) | **ADHD-C (N=32)** | **ADHD-I (N=47)** | **ADHD-H (N=1)** | **Statisticsc** |
| Age | 27.5 (5.2) | 26.3 (5.9) | 19.7 | *p*=0.352 |
| Sex (M/F) | 25/7 | 30/17 | 1/0 | *p*=0.319 |
| Handedness (R/L) | 24/8 | 38/9 | 1/0 | *p*=0.800 |
| FIQ | 108.8 (8.3) | 106,4 (11.5) | 115 | *p*=0.280 |
| VIQ | 108.8 (11.0) | 104.1 (11.1) | 103 | *p*=0.069 |
| PIQ | 110.3 (9.9) | 106.7 (19.1) | 128 | *p*=0.320 |
| *ADHD symptoms* |  |  |  |  |
| Inattentiona | 21.4 (3.8) | 18.6 (5.4) | 3 | *p*=0.011 |
| Hyperactivity/Impulsivitya | 17.6 (6.4) | 11.0 (5.0) | 13 | *p*<0.001 |
| Opposition-defiancea | 13.8 (5.3) | 9.9 (5.7) | 7 | *p*=0.003 |
| ASRS-A | 28.5 (3.4) | 26.4 (4.9) | 4 | *p*=0.037 |
| ASRS-B | 24.2 (4.0) | 17.4 (5.5) | 12 | *p*<0.001 |
| Mean frame-wise displacementb (mm) | 0.050 (0.025) | 0.048 (0.025) | 0.049 | *p*=0.737 |

**Supplementary Table 1b.** Demographics among attention-deficit hyperactivity (ADHD) subtypes (based on the childhood presentation of ADHD psychopathology)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mean (SD) | **ADHD-C (N=51)** | **ADHD-I (N=28)** | **ADHD-H (N=1)** | **Statisticsb** |
| Age | 27.1 (5.9) | 26.3 (5.2) | 19.7 | *p*=0.557 |
| Sex (M/F) | 39/12 | 16/12 | 1/0 | *p*=0.161 |
| Handedness (R/L) | 39/12 | 24/4 | 1/0 | *p*=0.564 |
| FIQ | 106.7 (9.6) | 108.5 (11.7) | 115 | *p*=0.461 |
| VIQ | 105.7 (11.8) | 106.5 (10.3) | 103 | *p*=0.744 |
| PIQ | 107.1 (17.0) | 110.0 (14.4) | 128 | *p*=0.443 |
| *ADHD symptoms* |  |  |  |  |
| Inattentiona | 19.7 (5.1) | 19.9 (4.8) | 3 | *p*=0.877 |
| Hyperactivity/Impulsivitya | 15.5 (6.4) | 10.3 (5.2) | 13 | *P*<0.001 |
| Opposition-defiancea | 11.8 (6.0) | 10.9 (5.8) | 7 | *p*=0.506 |
| ASRS-A | 27.2 (4.3) | 27.4 (4.8) | 4 | *p*=0.812 |
| ASRS-B | 21.7 (5.1) | 17.3 (6.5) | 12 | *p*=0.001 |
| Mean frame-wise displacementb (mm) | 0.049 (0.024) | 0.049 (0.026) | 0.049 | *p*=0.917 |

a Measured by the Swanson, Nolan, and Pelham, version IV (SNAP-IV) scale.

b Estimated by the Euclidian norm (enorm: square root of the sum of squares of the differences in motion derivatives), computed with AFNI's 1d\_tool.py.

c Statisitcal inference was only made from comparisons between ADHD-C and ADHD-I subgroups.

Abbreviation: -C=combined subtype; -I=inattentive subtype; -H=hyperactive-impulsive subtype; ASRS=Adult ADHD Self-Report Scale; FIQ=full-scale intelligence quotient; PIQ=performance intelligence quotient; VIQ=verbal intelligence quotient; M=male; F=female; R=right; L=left; SD=standard deviation.

**Supplementary Table 2.** Details of thenodes within the altered network of ADHD (network-based statistics, NBS)

|  |  |  |  |
| --- | --- | --- | --- |
|  | **MNI coordinates** | | |
| **Nodes** | x | y | z |
| DMN\_Frontal\_Sup\_R | 22 | 39 | 39 |
| DMN\_Occipital\_Mid\_L | -41 | -75 | 26 |
| DMN\_ParaHippocampal\_L | -26 | -40 | -8 |
| FPTC\_Frontal\_Mid\_L | -23 | 11 | 64 |
| FPTC\_Parietal\_Inf\_R | 44 | -53 | 47 |
| SN\_Precentral\_R | 42 | 0 | 47 |
| SN\_Insula\_L | -35 | 20 | 0 |
| SN\_Insula\_R | 36 | 22 | 3 |
| SN\_Cingulum\_Mid\_L | -1 | 15 | 44 |
| SN\_Frontal\_Mid\_R | 31 | 33 | 26 |
| SN\_Cingulum\_Mid\_L | 5 | 23 | 37 |
| COTC\_Frontal\_Sup\_L | -16 | -5 | 71 |
| COTC\_SupraMarginal\_R | 54 | -28 | 34 |
| COTC\_Rolandic\_Oper\_L | -45 | 0 | 9 |
| COTC\_Supp\_Motor\_Area\_R | 13 | -1 | 70 |
| COTC\_Insula\_R | 49 | 8 | -1 |
| COTC\_Temporal\_Pole\_Sup\_L | -51 | 8 | -2 |
| COTC\_Supp\_Motor\_Area\_R | 7 | 8 | 51 |
| COTC\_Insula\_R | 36 | 10 | 1 |
| COTC\_Cingulum\_Mid\_L | -5 | 18 | 34 |
| SN\_Frontal\_Mid\_R | 31 | 33 | 26 |
| SN\_Cingulum\_Mid\_L | 5 | 23 | 37 |
| COTC\_Frontal\_Sup\_L | -16 | -5 | 71 |
| COTC\_SupraMarginal\_R | 54 | -28 | 34 |
| COTC\_Rolandic\_Oper\_L | -45 | 0 | 9 |
| COTC\_Supp\_Motor\_Area\_R | 13 | -1 | 70 |
| COTC\_Insula\_R | 49 | 8 | -1 |
| COTC\_Temporal\_Pole\_Sup\_L | -51 | 8 | -2 |
| COTC\_Supp\_Motor\_Area\_R | 7 | 8 | 51 |
| COTC\_Insula\_R | 36 | 10 | 1 |
| COTC\_Cingulum\_Mid\_L | -5 | 18 | 34 |
| DAN\_Parietal\_Inf\_L | -33 | -46 | 47 |
| VAN\_Frontal\_Inf\_Tri\_L | -49 | 25 | -1 |
| subC\_Putamen\_L | -22 | 7 | -5 |
| subC\_Putamen\_R | 23 | 10 | 1 |
| subC\_Pallidum\_R | 15 | 5 | 7 |
| subC\_Thalamus\_R | 9 | -4 | 6 |
| subC\_Thalamus\_L | -2 | -13 | 12 |
| Vis\_Cuneus\_R | 15 | -77 | 31 |
| Vis\_Cuneus\_L | -16 | -77 | 34 |

Abbreviations: ADHD=attention-deficit hyperactivity disorder; MNI=Montreal Neurological Institute; DMN=default mode network; SN=salience network; COTC=cingulo-opercular network; FPTC=frontoparietal task control network; VAN=ventral attention network; DAN=dorsal attention network; SSM=somatosensorimotor network; Aud=auditory network; Vis=visual network; subC=subcortical; Supp=supplementary; R=right; L=left; Sup=superior; Inf=inferior; Mid=middle; Oper=opercular.

**Supplementary Table 3.** Average values of functional connectivity in the pairwise connections of interest (network-based statistics, NBS)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Pairs** | | **Control** | | | **ADHD** | | |
| **Network\_Region** | **Network\_Region** | **Mean** | **STE** | **STD** | **Mean** | **STE** | **STD** |
| DMN\_Occipital\_Mid\_L | SN\_Insula\_L | 0.211 | 0.022 | 0.241 | 0.339 | 0.029 | 0.263 |
| DMN\_Occipital\_Mid\_L | SN\_Insula\_R | 0.118 | 0.024 | 0.266 | 0.298 | 0.027 | 0.239 |
| DMN\_Frontal\_Sup\_R | SN\_Cingulum\_Mid\_L | 0.249 | 0.028 | 0.315 | 0.420 | 0.039 | 0.347 |
| DMN\_ParaHippocampal\_L | SN\_Cingulum\_Mid\_L | 0.108 | 0.023 | 0.257 | 0.249 | 0.032 | 0.286 |
| DMN\_Occipital\_Mid\_L | COTC\_Frontal\_Sup\_L | 0.238 | 0.028 | 0.311 | 0.420 | 0.036 | 0.324 |
| DMN\_Occipital\_Mid\_L | COTC\_SupraMarginal\_R | 0.273 | 0.023 | 0.258 | 0.414 | 0.031 | 0.279 |
| SN\_Precentral\_R | COTC\_Rolandic\_Oper\_L | 0.249 | 0.024 | 0.266 | 0.398 | 0.031 | 0.280 |
| DMN\_Occipital\_Mid\_L | COTC\_Supp\_Motor\_Area\_R | 0.114 | 0.027 | 0.294 | 0.278 | 0.036 | 0.321 |
| DMN\_Occipital\_Mid\_L | COTC\_Insula\_R | 0.057 | 0.025 | 0.279 | 0.210 | 0.033 | 0.293 |
| DMN\_Occipital\_Mid\_L | COTC\_Temporal\_Pole\_Sup\_L | 0.199 | 0.026 | 0.288 | 0.366 | 0.037 | 0.334 |
| DMN\_Occipital\_Mid\_L | COTC\_Supp\_Motor\_Area\_R | 0.170 | 0.025 | 0.276 | 0.308 | 0.031 | 0.275 |
| DMN\_ParaHippocampal\_L | COTC\_Supp\_Motor\_Area\_R | 0.156 | 0.023 | 0.251 | 0.284 | 0.026 | 0.229 |
| DMN\_Occipital\_Mid\_L | COTC\_Insula\_R | 0.180 | 0.023 | 0.253 | 0.310 | 0.029 | 0.255 |
| FPTC\_Frontal\_Mid\_L | COTC\_Cingulum\_Mid\_L | 0.472 | 0.029 | 0.317 | 0.630 | 0.030 | 0.268 |
| SN\_Cingulum\_Mid\_L | DAN\_Parietal\_Inf\_L | 0.332 | 0.023 | 0.257 | 0.494 | 0.032 | 0.290 |
| FPTC\_Parietal\_Inf\_R | VAN\_Frontal\_Inf\_Tri\_L | 0.073 | 0.025 | 0.273 | 0.212 | 0.029 | 0.264 |
| SN\_Insula\_R | VAN\_Frontal\_Inf\_Tri\_L | 0.129 | 0.027 | 0.303 | 0.283 | 0.029 | 0.260 |
| SN\_Frontal\_Mid\_R | VAN\_Frontal\_Inf\_Tri\_L | 0.062 | 0.022 | 0.244 | 0.193 | 0.026 | 0.228 |
| COTC\_Temporal\_Pole\_Sup\_L | VAN\_Frontal\_Inf\_Tri\_L | 0.362 | 0.028 | 0.307 | 0.541 | 0.032 | 0.288 |
| FPTC\_Frontal\_Mid\_L | subC\_Putamen\_L | 0.264 | 0.019 | 0.209 | 0.384 | 0.026 | 0.234 |
| DMN\_Occipital\_Mid\_L | subC\_Putamen\_R | 0.172 | 0.020 | 0.219 | 0.292 | 0.026 | 0.233 |
| FPTC\_Frontal\_Mid\_L | subC\_Putamen\_R | 0.283 | 0.020 | 0.219 | 0.419 | 0.022 | 0.199 |
| DMN\_ParaHippocampal\_L | subC\_Putamen\_L | 0.281 | 0.020 | 0.224 | 0.407 | 0.028 | 0.253 |
| DMN\_Frontal\_Sup\_R | subC\_Pallidum\_R | 0.212 | 0.020 | 0.221 | 0.347 | 0.033 | 0.291 |
| DMN\_Frontal\_Sup\_R | subC\_Thalamus\_R | 0.210 | 0.022 | 0.242 | 0.364 | 0.033 | 0.295 |
| DMN\_Frontal\_Sup\_R | subC\_Thalamus\_L | 0.163 | 0.024 | 0.266 | 0.323 | 0.033 | 0.292 |
| SN\_Insula\_R | SSM\_Postcentral\_L | 0.316 | 0.021 | 0.232 | 0.437 | 0.027 | 0.237 |
| COTC\_Rolandic\_Oper\_L | SSM\_Postcentral\_L | 0.451 | 0.027 | 0.299 | 0.596 | 0.027 | 0.243 |
| COTC\_Temporal\_Pole\_Sup\_L | SSM\_Postcentral\_L | 0.462 | 0.026 | 0.290 | 0.616 | 0.034 | 0.305 |
| COTC\_Supp\_Motor\_Area\_R | SSM\_Postcentral\_R | 0.295 | 0.028 | 0.308 | 0.449 | 0.031 | 0.279 |
| SN\_Insula\_R | SSM\_Precentral\_R | 0.332 | 0.024 | 0.261 | 0.472 | 0.026 | 0.234 |
| COTC\_Supp\_Motor\_Area\_R | SSM\_Precentral\_R | 0.364 | 0.027 | 0.300 | 0.537 | 0.029 | 0.260 |
| COTC\_Temporal\_Pole\_Sup\_L | SSM\_Postcentral\_L | 0.397 | 0.026 | 0.284 | 0.566 | 0.036 | 0.326 |
| SN\_Cingulum\_Mid\_L | SSM\_Insula\_R | 0.322 | 0.024 | 0.264 | 0.468 | 0.033 | 0.294 |
| COTC\_Supp\_Motor\_Area\_R | SSM\_Insula\_R | 0.319 | 0.025 | 0.272 | 0.478 | 0.032 | 0.290 |
| COTC\_Supp\_Motor\_Area\_R | SSM\_Postcentral\_L | 0.243 | 0.025 | 0.272 | 0.389 | 0.032 | 0.286 |
| COTC\_Supp\_Motor\_Area\_R | Aud\_Rolandic\_Oper\_L | 0.337 | 0.025 | 0.279 | 0.481 | 0.031 | 0.280 |
| SN\_Insula\_R | Vis\_Cuneus\_R | 0.212 | 0.022 | 0.247 | 0.337 | 0.028 | 0.251 |
| SN\_Cingulum\_Mid\_L | Vis\_Cuneus\_R | 0.257 | 0.023 | 0.255 | 0.396 | 0.030 | 0.267 |
| SN\_Insula\_R | Vis\_Cuneus\_L | 0.260 | 0.023 | 0.257 | 0.390 | 0.028 | 0.249 |
| COTC\_Supp\_Motor\_Area\_R | Vis\_Cuneus\_L | 0.268 | 0.022 | 0.247 | 0.409 | 0.025 | 0.224 |

Abbreviations: ADHD=attention-deficit hyperactivity disorder; STE=standard error; STD=standard deviation; DMN=default mode network; SN=salience network; COTC=cingulo-opercular network; FPTC=frontoparietal task control network; VAN=ventral attention network; DAN=dorsal attention network; SSM=somatosensorimotor network; Aud=auditory network; Vis=visual network; subC=subcortical; Supp=supplementary; R=right; L=left; Sup=superior; Inf=inferior; Mid=middle; Oper=opercular.

**Supplementary Table 4a.** The significant canonical correlation analysis (CCA) mode (*p*<0.05, family-wise error corrected) of the primary analysis.

|  |  |
| --- | --- |
| **CCA mode** | One |
| ***df1*** | 9 |
| ***df2*** | 180.25 |
| ***F*** | 2.03 |
| ***r*** | 0.430 |
| ***Wilk’s lambda*** | 0.7740 |
| ***Familywise error corrected p*** | 0.0367 |

**==================================================================**

**Supplementary Table 4b.** The significant CCA mode (*p*<0.05, family-wise error corrected) based on the 5 eigenvectors derived from the connectivity matrix.

|  |  |
| --- | --- |
| **CCA mode** | One |
| ***df1*** | 15 |
| ***df2*** | 199.17 |
| ***F*** | 1.61 |
| ***r*** | 0.446 |
| ***Wilk’s lambda*** | 0.7030 |
| ***Familywise error corrected p*** | 0.0491 |

**Supplementary Table 5.** Canonical correlation analysis (CCA) mode connectivity weight and associated interregional pairs

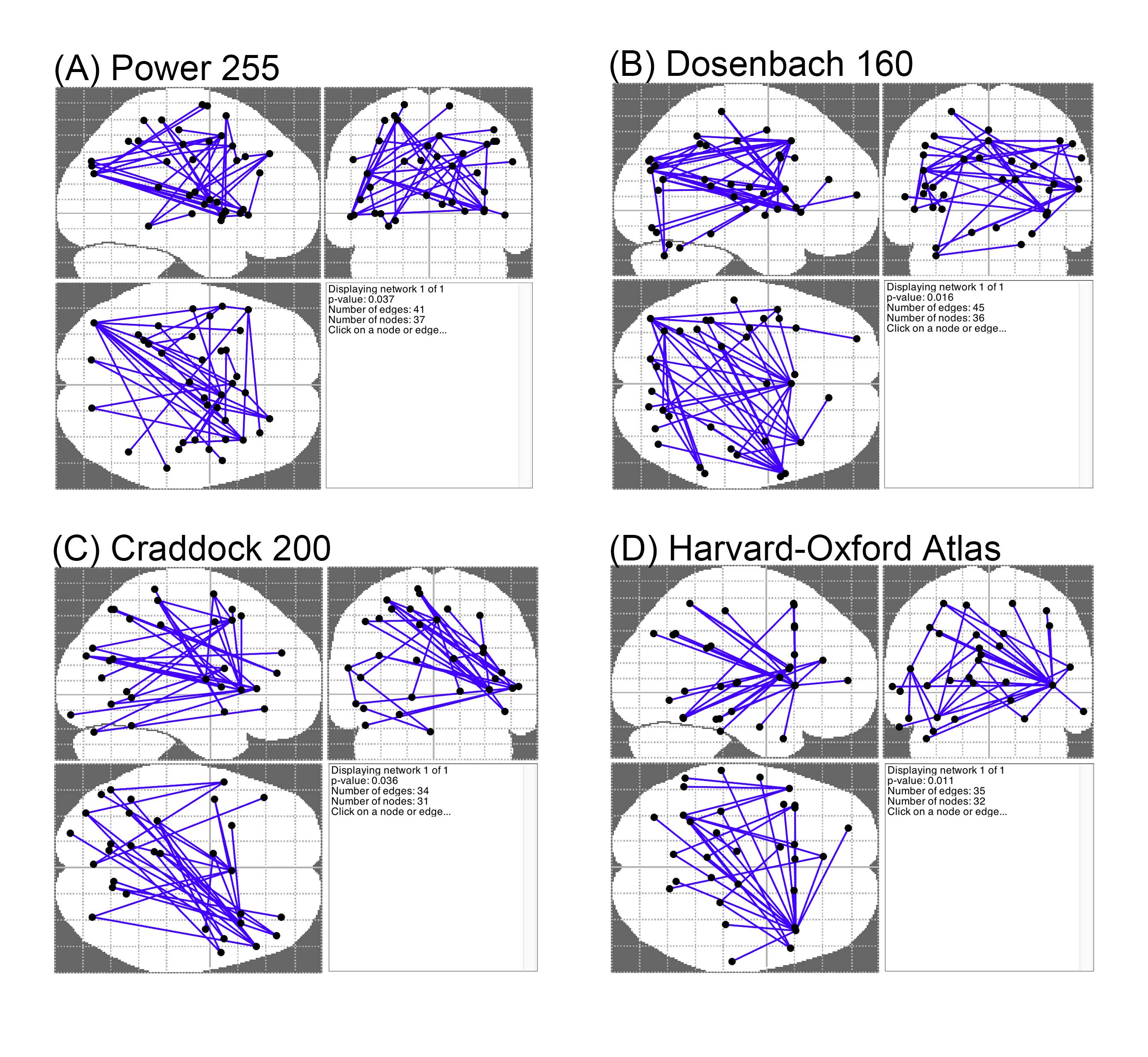
|  |  |  |
| --- | --- | --- |
| **Pairs** | | **CCA edge strength modulation** |
| **Network\_Region** | **Network\_Region** |
| DMN\_Occipital\_Mid\_L | SN\_Insula\_R | 0.549 |
| DMN\_Occipital\_Mid\_L | COTC\_Supp\_Motor\_Area\_R | 0.656 |
| DMN\_Occipital\_Mid\_L | COTC\_Insula\_R | 0.757 |
| DMN\_Occipital\_Mid\_L | COTC\_Temporal\_Pole\_Sup\_L | 0.755 |
| DMN\_Occipital\_Mid\_L | COTC\_Insula\_R | 0.623 |
| FPTC\_Frontal\_Mid\_L | COTC\_Cingulum\_Mid\_L | 0.568 |
| FPTC\_Frontal\_Mid\_L | subC\_Putamen\_L | 0.510 |
| DMN\_Occipital\_Mid\_L | subC\_Putamen\_R | 0.643 |
| FPTC\_Frontal\_Mid\_L | subC\_Putamen\_R | 0.510 |
| DMN\_Frontal\_Sup\_R | subC\_Pallidum\_R | 0.502 |
| DMN\_Frontal\_Sup\_R | subC\_Thalamus\_R | 0.500 |
| DMN\_Frontal\_Sup\_R | subC\_Thalamus\_L | 0.564 |

Abbreviations: DMN=default mode network; SN=salience network; COTC=cingulo-opercular network; FPTC=frontoparietal task control network; subC=subcortical; R=right; L=left; Sup=superior; Mid=middle; Supp=supplementary.

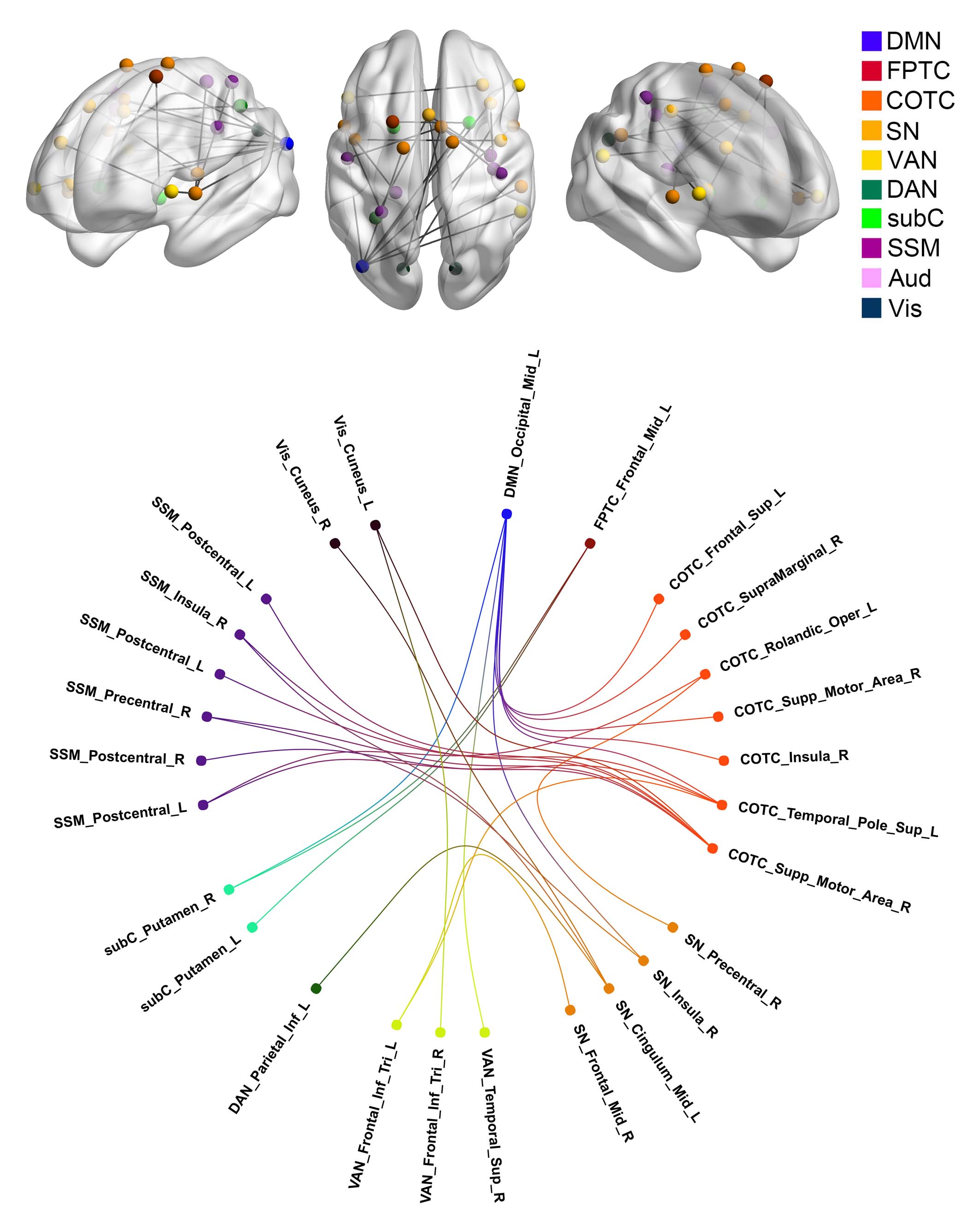
**Supplementary Table 6.** Validity indices of *k*-means clustering method (based on the feature vectors of individual participant’s weight derived from the connectivity and symptoms matrices of canonical correlation analysis)

|  |  |  |
| --- | --- | --- |
| ***k*-means clustering** | | |
|  | 2 clusters | 3 clusters |
| Jaccard similarity | 0.3385 | 0.5737 |
| Average silhouette width values | 0.4640 | 0.4861 |

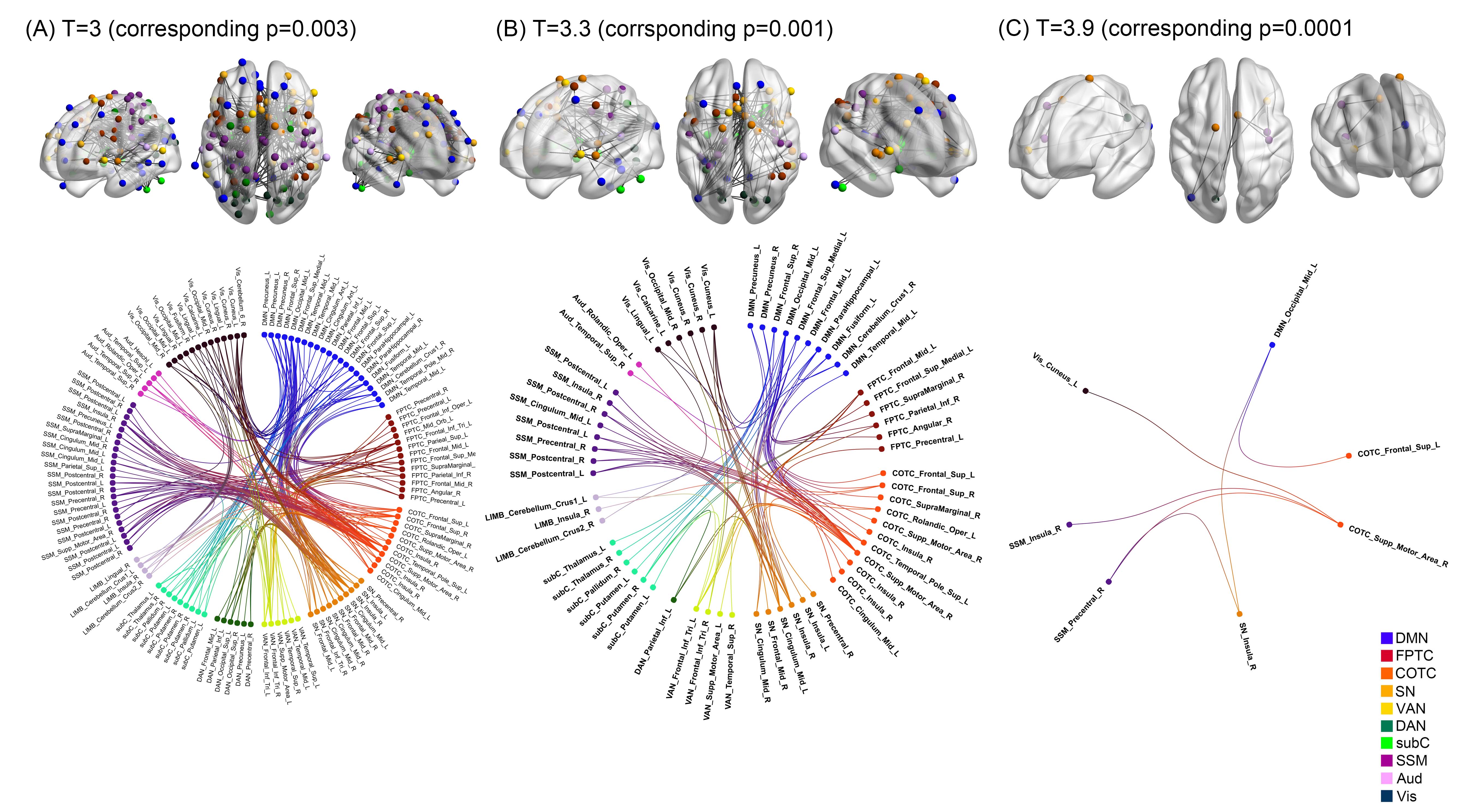
**3. Supplementary Figures**



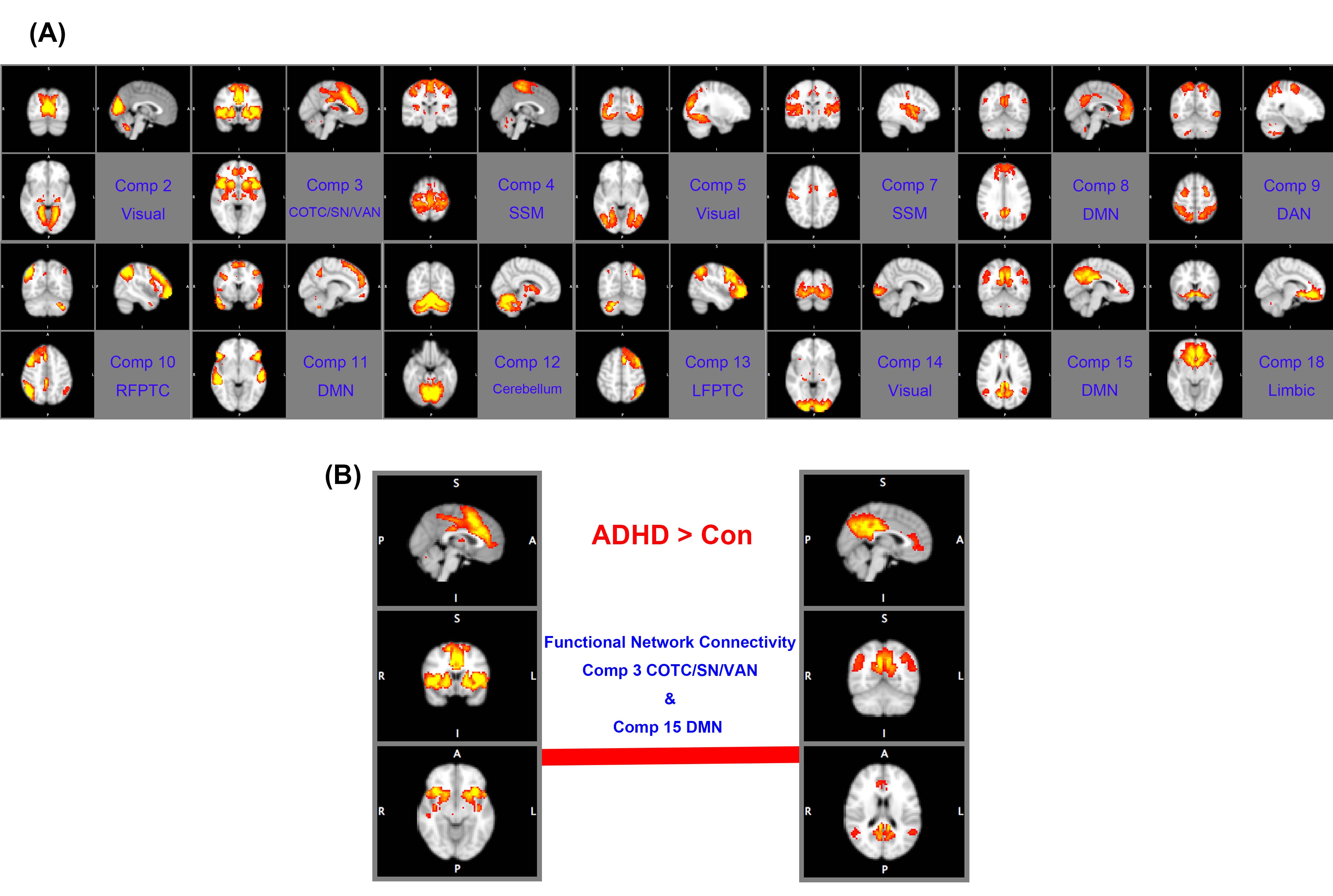
**Supplementary Figure 1.** *Changes in functional connectivity between adult ADHD and matched healthy controls across different brain parcellations.* The network-based statistic (NBS) showed stronger (generally stronger positive correlations, see Supplementary Table 3) functional connectivity in a single whole-brain network in ADHD compared to healthy controls. (A) 255 regions of interest parcellation (Power *et al.*, 2011). (B) 160 regions of interest parcellation (Dosenbach *et al.*, 2010). (C) 200 regions of interest parcellation (Craddock *et al.*, 2012). (D) the anatomical parcellation based on The Harvard-Oxford probabilistic cortical and subcortical atlases (www.fmrib.ox.ac.uk/fsl). Overall, the results obtained from different brain parcellations were similar.

**

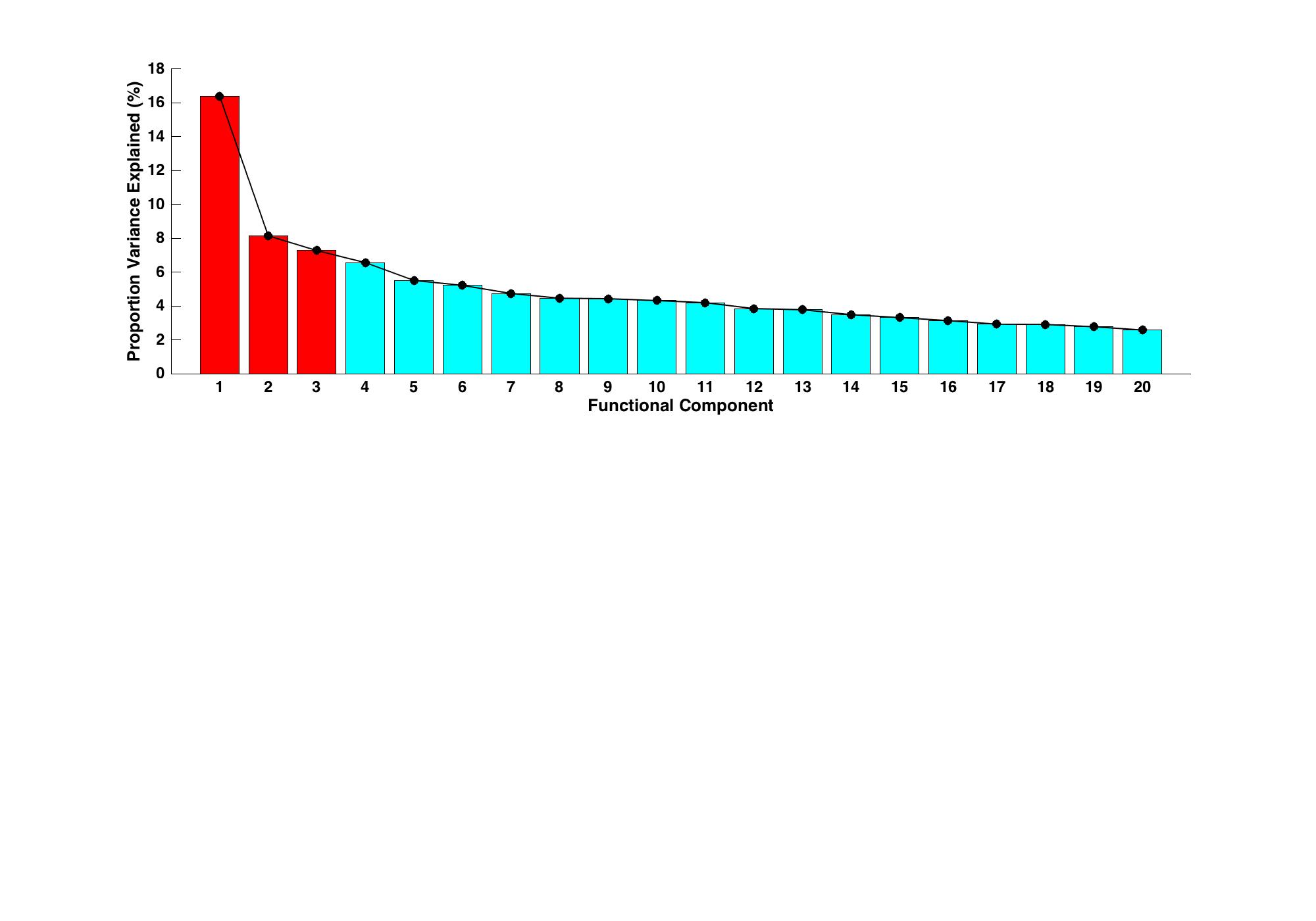
**Supplementary Figure 2.** *With additional adjusting for demographic features, group differences in inter-regional functional connectivity*. The network-based statistic (NBS) adjusting for gender/sex, levels of in-scanner head motion, and age identified a single network differentiating adult ADHD from healthy controls. This network was of largely the same pattern with the main analysis as shown in Figure 2: Adults with ADHD showed increased correlations between the DMN and frontoparietal network, the DMN and attention networks (including both salience/cingulo-opercular and dorsal attention components), the DMN and subcortical regions, the salience/cingulo-opercular network and sensory-motor and visual network, as well as the salience/cingulo-opercular network and dorsal attention alongside frontoparietal networks.

**

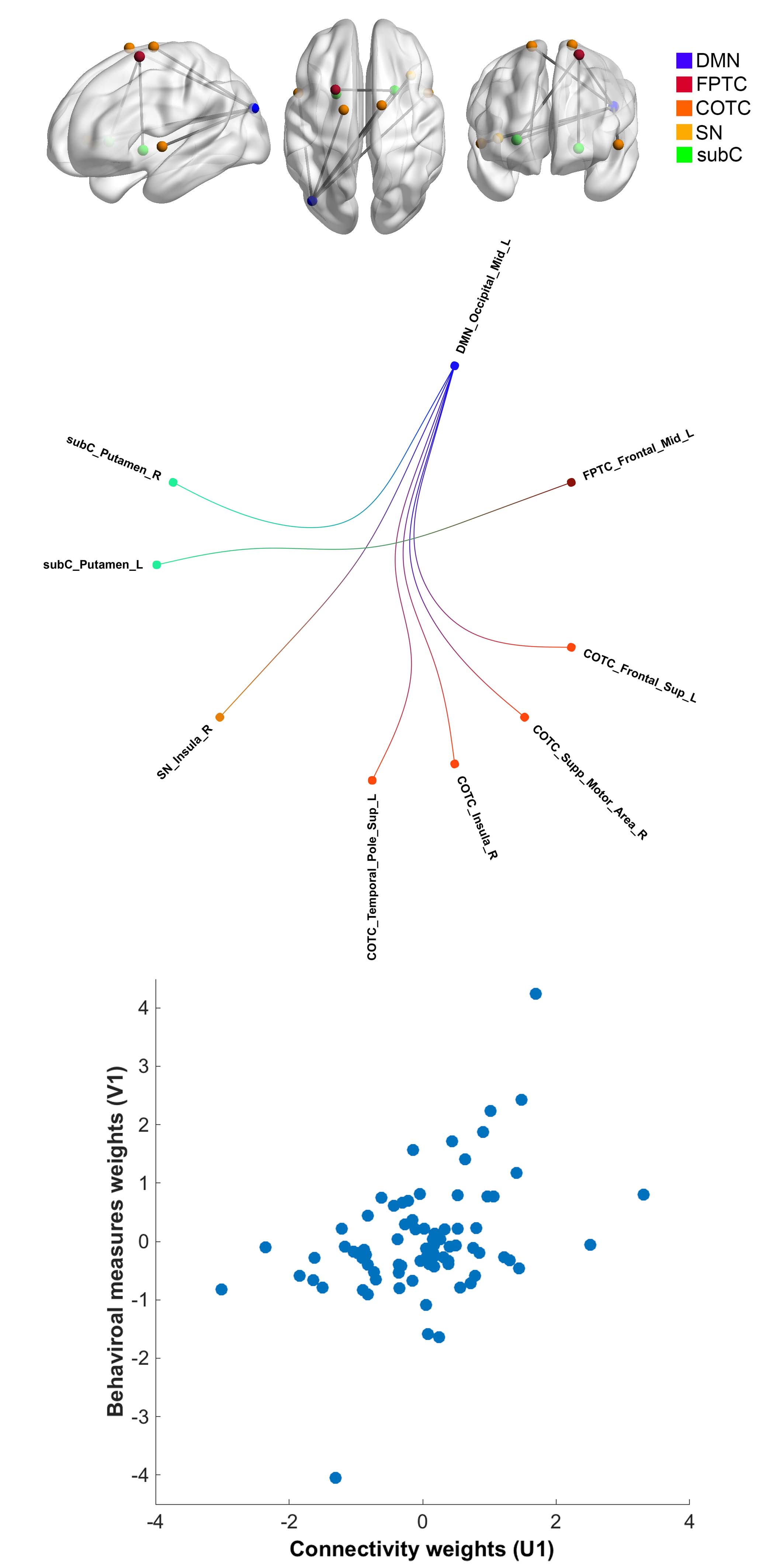
**Supplementary Figure 3.** *Group differences in functional connectivity based on different height thresholds in the network-based statistic (NBS).* Across different t-statistics height thresholds (namely t=3 corresponding to uncorrected p=0.003; t=3.3 corresponding to uncorrected p=0.001; t=3.9 corresponding to uncorrected p=0.0001), the NBS consistently identified the similar patterns of hyperconnectivity between the DMN and attention and cognitive control networks, the DMN and subcortical regions, and the salience/cingulo-opercular network and sensory processing networks. These results were in line with the main discovered connectivity differences based on t=3.5 corresponding to uncorrected p=0.0005.



**Supplementary Figure 4.** *Independent component analysis (ICA) on neuroimaging data.* (A) Based on the group ICA, we identified 20 spatial components. The topology of 14 components related to recognized functional brain networks (Yeo *et al.*, 2011; Cocchi *et al.*, 2012). These 14 components were used for the confirmatory functional network connectivity analysis (see text) (Jafri *et al.*, 2008). (B) Results from the functional network connectivity analysis are presented. Relative to the controls, adults with ADHD exhibited a significantly increased positive interaction between the default-mode and cingulo-opercular/salience networks (false discovery rate corrected *q*=0.044).



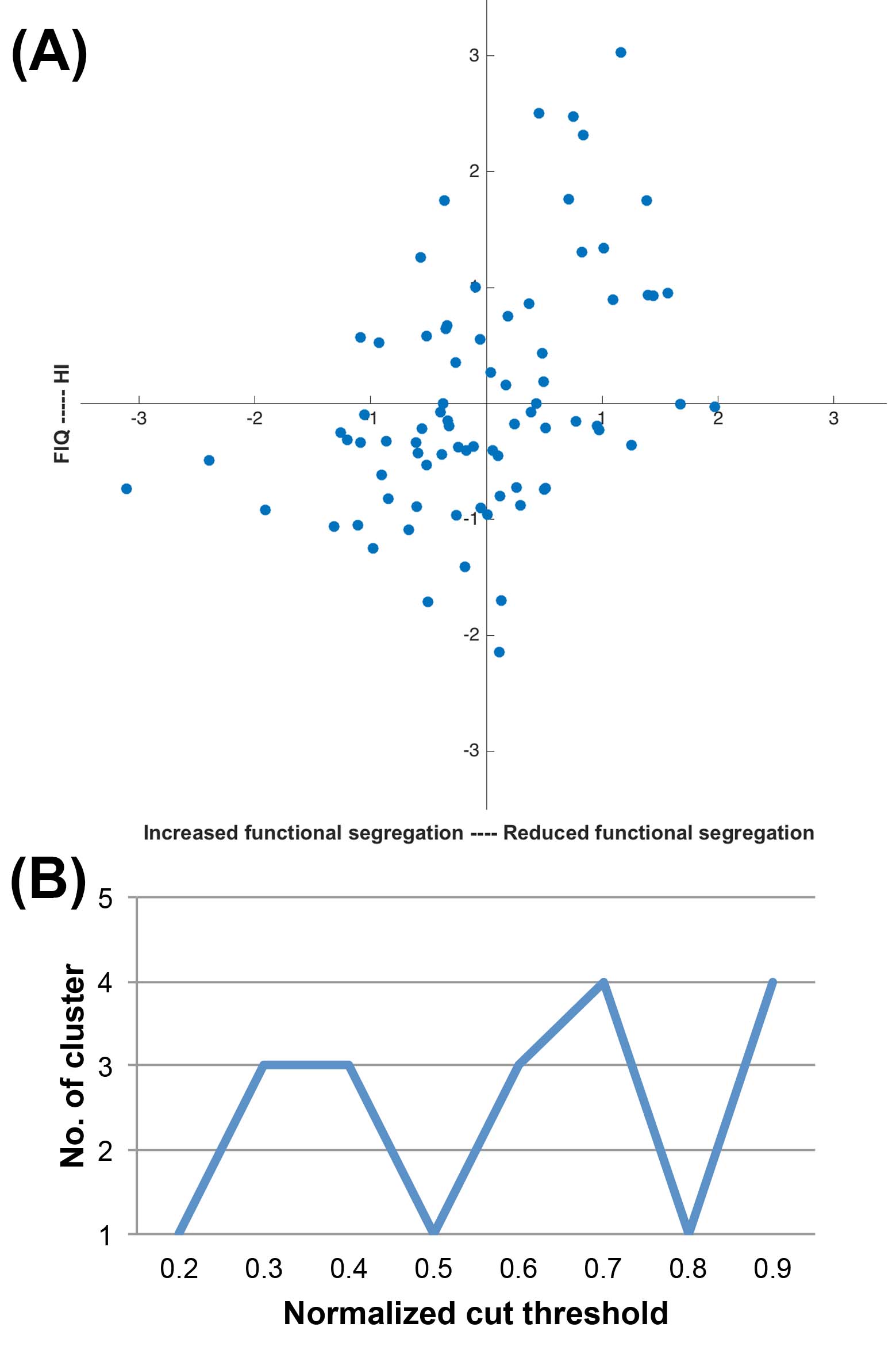
**Supplementary Figure 5.** *The proportion of variance explained by the eigenvectors defined by a principal component analysis on functional connectivity differences between ADHD and controls (derived from the network-based statistics).* The three eigenvectors (red) used in the canonical correlation analysis (CCA, see text) explained 31.83% of the total variance in between-groups connectivity. Including two extra eigenvectors allows to explain 43.90% of the variance. CCA based on both three and five eigenvectors yielded a similarly significant CCA mode.



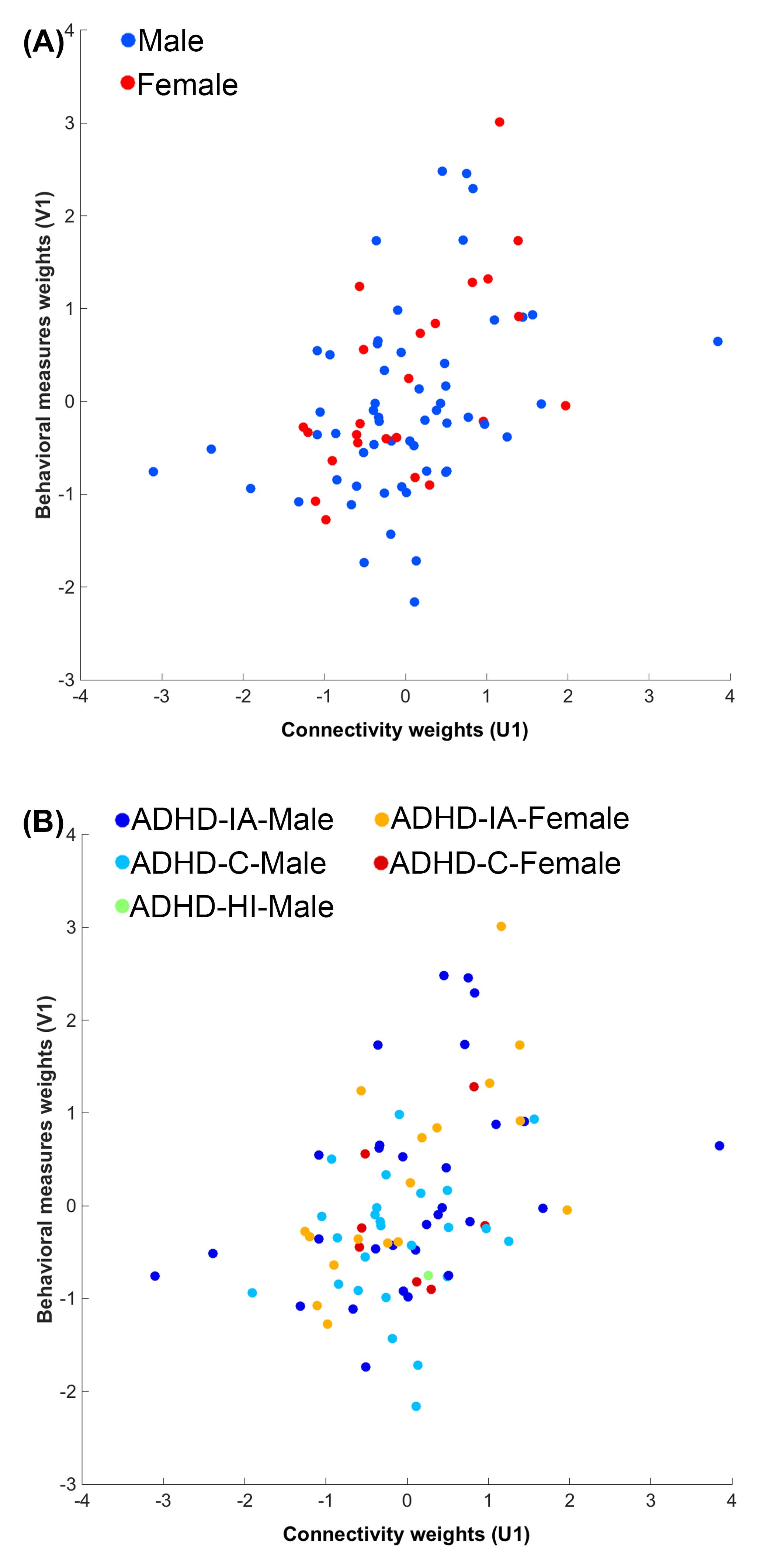
**Supplementary Figure 6.** *Supplementary**canonical correlation analysis (CCA) based on the altered functional connectivity identified by the supplementary network-based statistic (NBS) adjusting for gender/sex, levels of in-scanner head motion, and age.* This supplementary CCA yielded one significant mode, similar to the main result (Figure 3), which linked the brain connectivity and clinical symptoms-intelligence. The functional connections expressing the strongest positive associations in this mode from the supplementary analysis also implicated connectivity between the DMN and cingulo-opercular, as well as the DMN and subcortical regions.



**Supplementary Figure 7.** *The gap statistic for the**k-means clustering method (based on the feature vectors of individual participant’s weight derived from the connectivity and symptoms matrices of canonical correlation analysis).* The gap statistic estimates the optimal number of clusters by searching the local maximum of the graph, then selecting the smallest *k* within one standard error (as indicated by the bars in the figure) of the local max [Gap(*k*) ≥ Gap(*k+1*) – *SEk+1*]. Based on the gap statistic, the suggested optimal number of clusters was one.



**Supplementary Figure 8.** *Test for**ADHD categorical biotypes*. (A) *K*-means analysis failed to reveal valid clusters based on the individual associations between functional connectivity and behavior. The absence of clear clusters in the data is evident from visual inspection of the figure. (B) The number (No.) of clusters detected by the multi-view spectral clustering algorithm changed as a function of the preset cut threshold, indicating that no stable decomposition was achievable. Overall, results from these analyses provide compelling evidence for the absence of non-overlapping clusters in the data. FIQ=full-scale IQ; HI=hyperactivity-impulsivity.

****

**Supplementary Figure 9.** *(A) Males and females with ADHD, (B) regardless of the clinical subtypes, were distributed evenly along the one-dimensional axis identified by the main CCA.*

**4. References for Supplementary Materials**

**Abdi, H. and Williams, L. J.** (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics* **2**, 433-459.

**Asherson, P., Buitelaar, J., Faraone, S. V. and Rohde, L. A.** (2016). Adult attention-deficit hyperactivity disorder: key conceptual issues. *Lancet Psychiatry* **3**, 568-578.

**Beckmann, C. F. and Smith, S. M.** (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging* **23**, 137-152.

**Benjamini, Y. and Hochberg, Y.** (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300.

**Chang, L. R., Chiu, Y. N., Wu, Y. Y. and Gau, S. S.** (2013). Father's parenting and father-child relationship among children and adolescents with attention-deficit/hyperactivity disorder. *Comprehensive Psychiatry* **54**, 128-140.

**Chao, C. Y., Gau, S. S., Mao, W. C., Shyu, J. F., Chen, Y. C. and Yeh, C. B.** (2008). Relationship of attention-deficit-hyperactivity disorder symptoms, depressive/anxiety symptoms, and life quality in young men. *Psychiatry and Clinical Neurosciences* **62**, 421-426.

**Chen, H., Li, K., Zhu, D., Jiang, X., Yuan, Y., Lv, P., Zhang, T., Guo, L., Shen, D. and Liu, T.** (2013). Inferring group-wise consistent multimodal brain networks via multi-view spectral clustering. *IEEE Transactions on Medical Imaging* **32**, 1576-1586.

**Cocchi, L., Harrison, B. J., Pujol, J., Harding, I. H., Fornito, A., Pantelis, C. and Yucel, M.** (2012). Functional alterations of large-scale brain networks related to cognitive control in obsessive-compulsive disorder. *Human Brain Mapping* **33**, 1089-1106.

**Craddock, R. C., James, G. A., Holtzheimer, P. E., 3rd, Hu, X. P. and Mayberg, H. S.** (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping* **33**, 1914-1928.

**DiStefano, C., Zhu, M. and Mindrila, D.** (2009). Understanding and using factor scores: considerations for the applied researcher. *Practical Assessment, Research & Evaluation* **14**, 1-11.

**Dolnicar, S.** (2002). A review of unquestioned standards in using cluster analysis for data-driven market segmentation. In *CD Conference Proceedings of the Australian and New Zealand Marketing Academy Conference 2002 (ANZMAC 2002)*: Deakin University, Melbourne.

**Dosenbach, N. U., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., Nelson, S. M., Wig, G. S., Vogel, A. C., Lessov-Schlaggar, C. N., Barnes, K. A., Dubis, J. W., Feczko, E., Coalson, R. S., Pruett, J. R., Jr., Barch, D. M., Petersen, S. E. and Schlaggar, B. L.** (2010). Prediction of individual brain maturity using fMRI. *Science* **329**, 1358-1361.

**Gau, S. F. and Soong, W. T.** (1999). Psychiatric comorbidity of adolescents with sleep terrors or sleepwalking: a case-control study. *Australian & New Zealand Journal of Psychiatry* **33**, 734-739.

**Gau, S. S., Chong, M. Y., Chen, T. H. and Cheng, A. T.** (2005). A 3-year panel study of mental disorders among adolescents in Taiwan. *The American Journal of Psychiatry* **162**, 1344-1350.

**Gau, S. S., Kessler, R. C., Tseng, W. L., Wu, Y. Y., Chiu, Y. N., Yeh, C. B. and Hwu, H. G.** (2007). Association between sleep problems and symptoms of attention-deficit/hyperactivity disorder in young adults. *Sleep* **30**, 195-201.

**Gau, S. S., Shang, C. Y., Liu, S. K., Lin, C. H., Swanson, J. M., Liu, Y. C. and Tu, C. L.** (2008). Psychometric properties of the Chinese version of the Swanson, Nolan, and Pelham, version IV scale - parent form. *International Journal of Methods in Psychiatric Research* **17**, 35-44.

**Hennig, C.** (2008). Dissolution point and isolation robustness: robustness criteria for general cluster analysis methods. *Journal of Multivariate Analysis* **99**, 1154-1176.

**Huettel, S. A., Song, A. W. and McCarthy, G.** (2008). *Functional Magnetic Resonance Imaging.* Sinauer Associates: Sunderland.

**Hyvarinen, A.** (1999a). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans Neural Netw* **10**, 626-634.

**Hyvarinen, A.** (1999b). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks* **10**, 626-634.

**Jafri, M. J., Pearlson, G. D., Stevens, M. and Calhoun, V. D.** (2008). A method for functional network connectivity among spatially independent resting-state components in schizophrenia. *NeuroImage* **39**, 1666-1681.

**Kononenko, I. and Kukar, M.** (2007). *Machine Learning and Data Mining: Introduction to Principles and Algorithms.* Woodhead Publishing.

**Kumar, A. and Daumé, H.** (2011). A co-training approach for multi-view spectral clustering. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*pp. 393-400.

**Kundu, P., Brenowitz, N. D., Voon, V., Worbe, Y., Vertes, P. E., Inati, S. J., Saad, Z. S., Bandettini, P. A. and Bullmore, E. T.** (2013). Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *Proceedings of the National Academy of Sciences* **110**, 16187-16192.

**Kundu, P., Inati, S. J., Evans, J. W., Luh, W. M. and Bandettini, P. A.** (2012). Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *NeuroImage* **60**, 1759-1770.

**Lv, J., Iraji, A., Ge, F., Zhao, S., Hu, X., Zhang, T., Han, J., Guo, L., Kou, Z. and Liu, T.** (2016). Temporal concatenated sparse coding of resting state fMRI data reveal network interaction changes in mTBI. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*pp. 46-54. Springer.

**Ni, H. C., Lin, Y. J., Gau, S. S., Huang, H. C. and Yang, L. K.** (2017). An open-label, randomized trial of methylphenidate and atomoxetine treatment in adults With ADHD. *Journal of Attention Disorders* **21**, 27-39.

**Ni, H. C., Shang, C. Y., Gau, S. S., Lin, Y. J., Huang, H. C. and Yang, L. K.** (2013). A head-to-head randomized clinical trial of methylphenidate and atomoxetine treatment for executive function in adults with attention-deficit hyperactivity disorder. *The International Journal of Neuropsychopharmacology* **16**, 1959-1973.

**Orvaschel, H., Puig-Antich, J., Chambers, W., Tabrizi, M. A. and Johnson, R.** (1982). Retrospective assessment of prepubertal major depression with the Kiddie-SADS-e. *Journal of the American Academy of Child Psychiatry* **21**, 392-397.

**Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L. and Petersen, S. E.** (2011). Functional network organization of the human brain. *Neuron* **72**, 665-678.

**Shi, J. and Malik, J.** (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**, 888-905.

**Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R. and Beckmann, C. F.** (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences* **106**, 13040-13045.

**Smith, S. M., Hyvarinen, A., Varoquaux, G., Miller, K. L. and Beckmann, C. F.** (2014). Group-PCA for very large fMRI datasets. *NeuroImage* **101**, 738-749.

**Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E., Glasser, M. F., Ugurbil, K., Barch, D. M., Van Essen, D. C. and Miller, K. L.** (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature Neuroscience* **18**, 1565-1567.

**Swanson, J. M., Kraemer, H. C., Hinshaw, S. P., Arnold, L. E., Conners, C. K., Abikoff, H. B., Clevenger, W., Davies, M., Elliott, G. R., Greenhill, L. L., Hechtman, L., Hoza, B., Jensen, P. S., March, J. S., Newcorn, J. H., Owens, E. B., Pelham, W. E., Schiller, E., Severe, J. B., Simpson, S., Vitiello, B., Wells, K., Wigal, T. and Wu, M.** (2001). Clinical relevance of the primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment. *Journal of the American Academy of Child and Adolescent Psychiatry* **40**, 168-179.

**Takahashi, M., Goto, T., Takita, Y., Chung, S. K., Wang, Y. and Gau, S. S.** (2014). Open-label, dose-titration tolerability study of atomoxetine hydrochloride in Korean, Chinese, and Taiwanese adults with attention-deficit/hyperactivity disorder. *Asia-Pacific psychiatry : official journal of the Pacific Rim College of Psychiatrists* **6**, 62-70.

**Tao, H., Hou, C. and Yi, D.** (2014). Multiple-view spectral embedded clustering using a co-training approach. In *Computer Engineering and Networking*pp. 979-987. Springer.

**Tibshirani, R., Walther, G. and Hastie, T.** (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **63**, 411-423.

**Venkataraman, A., Van Dijk, K. R., Buckner, R. L. and Golland, P.** (2009). Exploring functional connectivity in fMRI via clustering. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing* **2009**, 441-444.

**Yang, H. N., Tai, Y. M., Yang, L. K. and Gau, S. S.** (2013). Prediction of childhood ADHD symptoms to quality of life in young adults: adult ADHD and anxiety/depression as mediators. *Research in Developmental Disabilities* **34**, 3168-3181.

**Yeh, C. B., Gau, S. S., Kessler, R. C. and Wu, Y. Y.** (2008). Psychometric properties of the Chinese version of the adult ADHD Self-report Scale. *International Journal of Methods in Psychiatric Research* **17**, 45-54.

**Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zollei, L., Polimeni, J. R., Fischl, B., Liu, H. and Buckner, R. L.** (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology* **106**, 1125-1165.