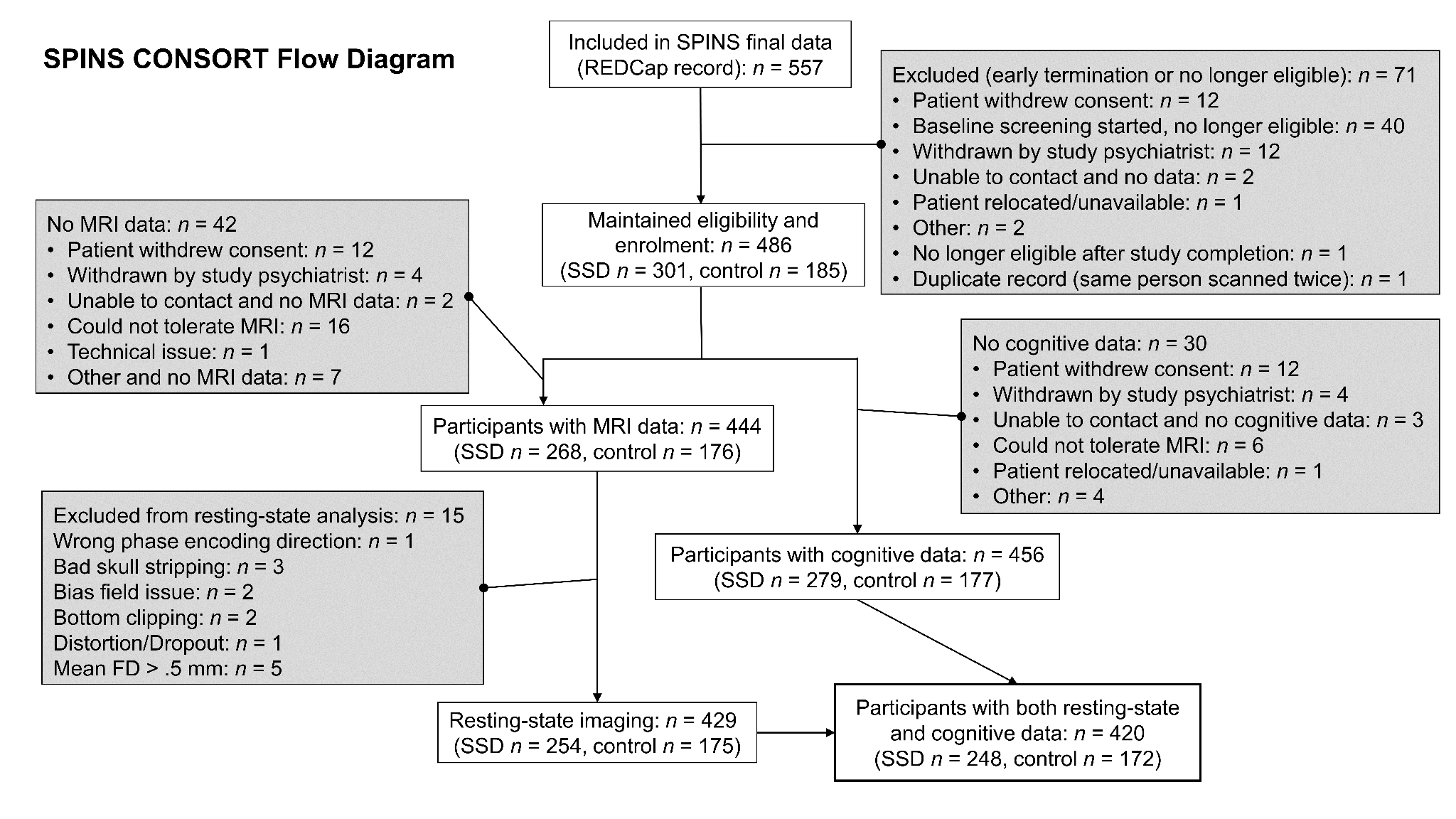
**Supplementary materials**

**Methods**

***Participants***

The following figure illustrates the inclusion/exclusion criteria at each stage of data analysis.



**Figure S1.** The data from SPINS recruited 557 participants with 486 eligible participants (301 for SSD and 185 for controls). We excluded data for statistical analysis based on quality control (QC) criteria that include screenings of the structural image (i.e., Dashboard QC), the functional magnetic resonance imaging (fMRI) preprocessing outputs (i.e., fMRIPrep/Ciftify QC), and excessive motion captured by framewise displacement (FD). The final sample includes 248 participants with SSDs and 172 healthy controls who have both complete cognitive and resting-state fMRI data.

***MRI Data Acquisition***

MRI data was collected across 6 scanners, including a General Electric Discovery (*N* = 135) and Siemens Prisma (*N* = 30) at CAMH, a General Electric Signa (*N* = 42) and Siemens Prisma (*N* = 98) at ZHH, and a Siemens Tim Trio (*N* = 66) and Siemens Prisma (*N* = 79) at MPRC. To ensure sequence stability over time and minimize inter-site variance, standardized operating procedures were used along with weekly phantom scans. The study also provided objective evidence for inter-site stability (1–3) . Due to scanner differences there were slight variations in the parameters for the T1 MRIs: CMH and ZHH used a BRAVO sequence with TR=6.7/6.4 ms, TE=3/2.8 ms, flip angle=8°, field of view=230 mm, in plane resolution=0.9 mm2, slice thickness=0.9 mm; CMP, MRC, MRP, and ZHP used an MPRAGE sequence with TR=2300 ms, TE=2.9 ms, flip angle=9°, field of view=230 mm, in plane resolution=0.9 mm2, slice thickness=0.9 mm).The Resting State (RS) scan was also part of a longer multimodal MRI protocol previously described (4).

***MRI Preprocessing***

The fMRI data were preprocessed using fMRIPrep 1.5.8 (5) and Nipype 1.4.1 (6). Anatomical T1-weighted images were corrected for intensity non-uniformity and skull-stripped using ANTs 2.2.0, and brain tissue segmentation was performed by FSL 5.0.9 (7). Brain surfaces were reconstructed using FreeSurfer 6.0.1 (8), For each fMRI run, ANTs (9) was used to perform fieldmap-less distortion correction, and the Freesurfer’s boundary-based registration, with six degrees of freedom, was performed for co-registration of the functional data to the corresponding T1-weighted image. Slice-timing correction and motion correction were performed using MCFLIRT [(FSL 5.0.9) (10)](https://www.zotero.org/google-docs/?XYXBsv).

Following fMRIPrep, the ciftify workflows (11) version 2.3.1 were used to convert the freesurfer reconstructed surfaces to gifti and cifti file formats. The cortical surfaces were realigned to the HCP fsLR templates (12), using sulcal depth using the MSM algorithm [(MSMSulc) (13)](https://www.zotero.org/google-docs/?8OpMDU) and resampled to 32k vertices per hemisphere, and the freesurfer subcortical segmentation was used to define the participants 32k subcortical atlas greyordinates. The functional data was projected to the 32k surface coordinates using a ribbon constrained method that excludes outlier voxels, with methods similar to those employed by the HCP Pipelines (12).

We dropped the first three TRs for each scan, and performed spatial smoothing with a 2 mm full width at half maximum Gaussian kernel. ‘ciftify\_clean\_img’ was then used to detrend and band-pass filter (0.01-0.1 Hz) the signals and perform nuisance regression on the data. The nuisance regression model included six head motion correction parameters, mean white matter signal, mean cerebral spinal fluid signal, mean global signal and the square, the derivative, and the squared derivative for each of these regressors (generated by fMRIPrep). We regressed out the mean global signal to remove the dominating global effects.

***Statistical Analysis***

For any participants missing one cognitive measure, data were imputed with the rest of the behavioral variables using the *mice* package in R. To examine group differences in cognition, we first examined the homoscedasticity between groups by *F*-tests and confirmed the normality of each group by Shapiro-Wilk’s test. Next, we examined group differences between SSDs and controls before imputing missing data. We performed the two-sample equal variances *t*-test if the measure showed homoscedasticity between groups, and we performed the Welch's unequal variances *t*-test if the groups are heteroscedastic. If either group did not pass the normality test, we performed a non-parametric bootstrap *t*-test, which has no normality assumption of the data.

***Partial Least Squares Correlation (PLSC)***

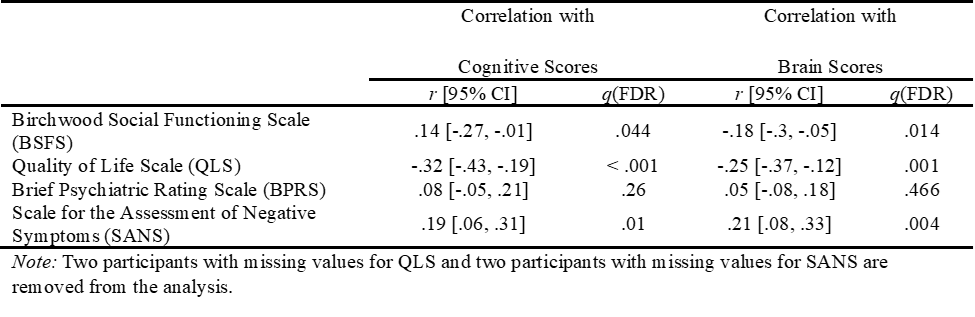
Partial least squares correlation (PLSC) (14) is a component-based method that is used to examine the association between two sets of variables measuring the same participants. PLSC decomposes the cross product between the two variable sets. In our analysis, with the variables centered and normalized to have sums of squares of 1s, such cross product gives scaled correlations between the two sets of variables. PLSC then decomposes such cross product matrix into uncorrelated *dimensions* that best capture different components of the correlation pattern. In PLSC, each dimension is composed of 1) two sets of *latent variables*, which represent the participants on this dimension, with respect to the two sets of variables (i.e., behavioral and gradients), and 2) two sets of variable *loadings*, which describe how they contribute to this dimension.

Formally, latent variables are new variables obtained by linear combinations (weighted sums) of the original variables. Each dimension of PLSC includes one pair of latent variables—one obtained from each data table—and the coefficients used to compute these linear combinations are the variable loadings. In PLSC, the first dimension is extracted such that the pair of latent variables are as similar as possible, as measured by their covariance; this covariance is quantified by the first *singular value* of PLSC. Subsequently, the latent variables of the second dimension are obtained from the residual data of the first dimension and explain the maximum covariance from the residual data with this covariance stored as the second singular value. Subsequent dimensions are obtained in the same way. Together, PLSC dimensions explain the covariance of the data tables in a descending order and are orthogonal to each other. Given such orthogonality, these dimensions explain independent sources of covariance, and the squared singular values (i.e., the *eigenvalues*, denoted by *λ*) are thus additive. Specifically, they add up to the total squared scaled correlation of the two data tables, and the ratio of each eigenvalue with respect to this total quantifies the proportion of scaled correlation explained by each dimension (denoted by *τ*). As *τ* is similar to the idea of a proportion of variance explained as measured by *η*2, *τ* can be interpreted as the effect size of a PLSC dimension.

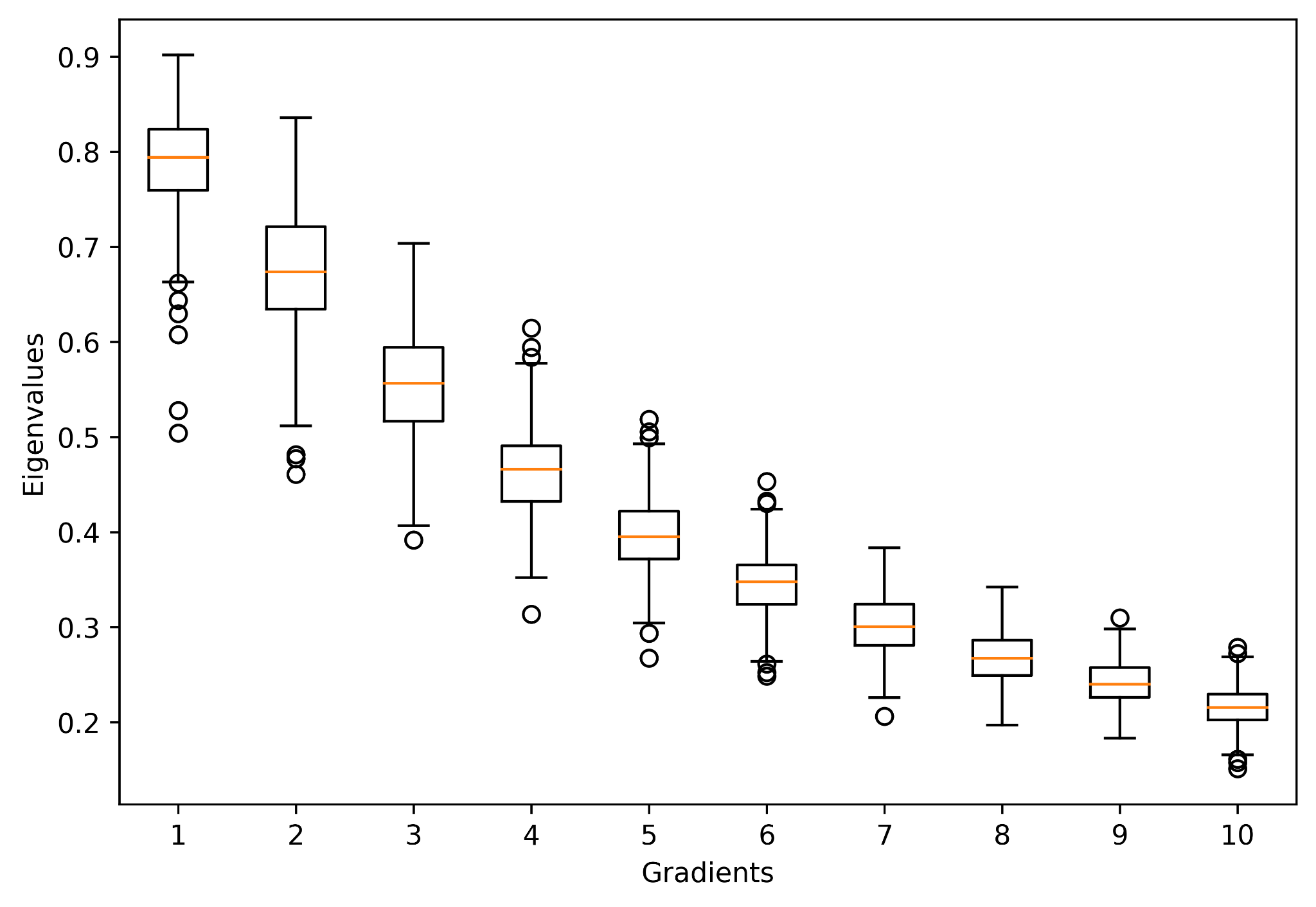
**Identify reliable dimensions.** To identify reliable dimensions, we performed a permutation test on the singular values to determine if the covariance between the corresponding pair of latent variables are reliably larger than 0. In the permutation test, we first permuted the participants within each variable of both tables such that the relationships between them were null. The permuted tables were then analyzed by PLSC to extract the first singular value. This procedure was repeated 1000 times to generate the null distribution of the first singular value. For the second dimension, the first dimension was first regressed out from the data before the 1000 iterations of permutation and PLSC to obtain the null distribution of the second singular value; a similar procedure was used for subsequent dimension. Just like in the null hypothesis testing, we then compared the observed values to their corresponding null distribution and obtained their *p* value as the probability associated with each observed value under their null. A singular value with *p* < .05 indicates a reliable dimension (*α* = .05).

**Identify important variables and reliable loadings.** To identify important variables for a given dimension, we quantified the *contribution* of each variable by computing the ratio of its squared loading to the eigenvalue. Because each variable was first normalised to have a sum of squares of 1, it originally contributed 1/*J* of variance to the total variance of the data table (with *J* being the number of variables). A contribution larger than 1/*J* therefore marks an important variable as it contributes more than average to the variance of a given dimension.

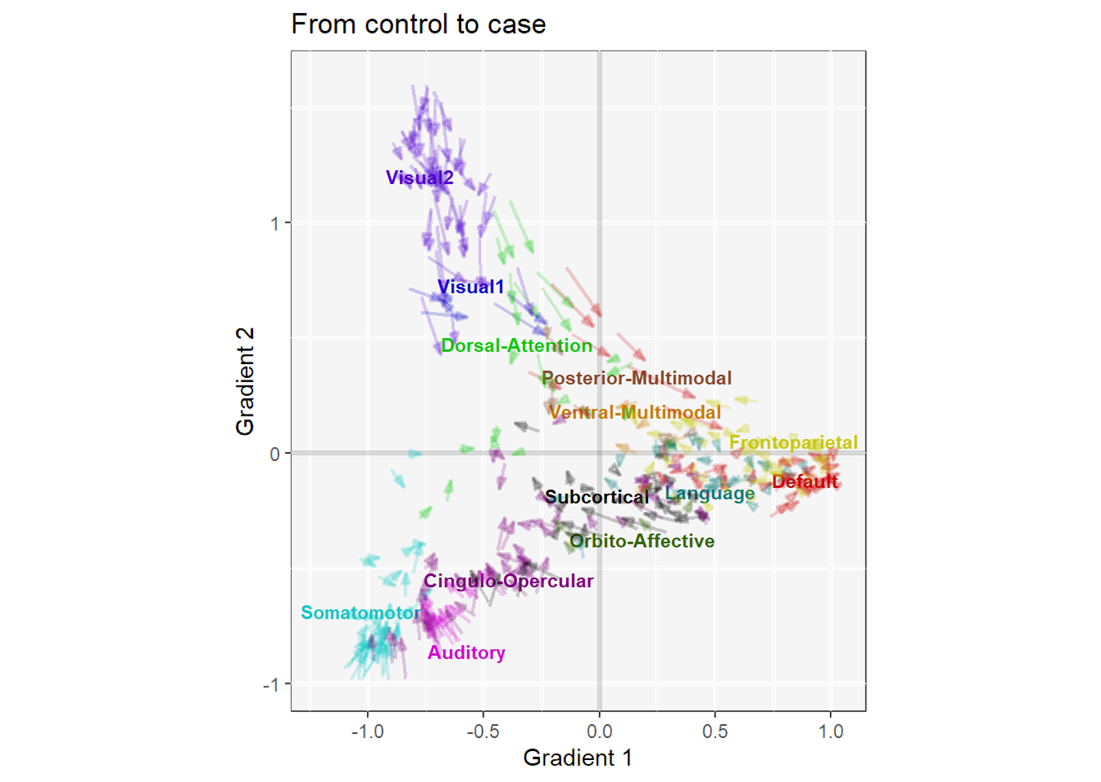
To identify reliable loadings, we used bootstrap tests to estimate the stability of the loadings. The bootstrap procedure generates a sampling distribution of the given measure (here, the loadings) assuming that the current sample is the best approximate of the population. Therefore, the bootstrap procedure generates subsamples from the original data set to estimate how the given measure varies. In the bootstrap procedure of PLSC, the participants were resampled with replacement to reconstruct the two resampled data tables. These two tables were centred by the original variable means and normalised by the original variable sums of squares then analysed by a PLSC to generate the loadings. This procedure was repeated 1000 times to generate the bootstrapped sampling distribution of all loadings. The reliability was then quantified by the *bootstrap ratio* (BR), which was computed by dividing each loading by the standard deviation of its bootstrapped sample distribution. Mathematically, BR is a *Z*-approximate that indicates whether the observed loading is reliably different from 0. Therefore, just like the Z test, a BR of 1.96 was used as the critical value to indicate reliability at *α* of .05. In this paper, regarding the excessive number of tests, we used a more stringent critical value of 2.88 for two-tail *Z* tests at *α* < .005.

****

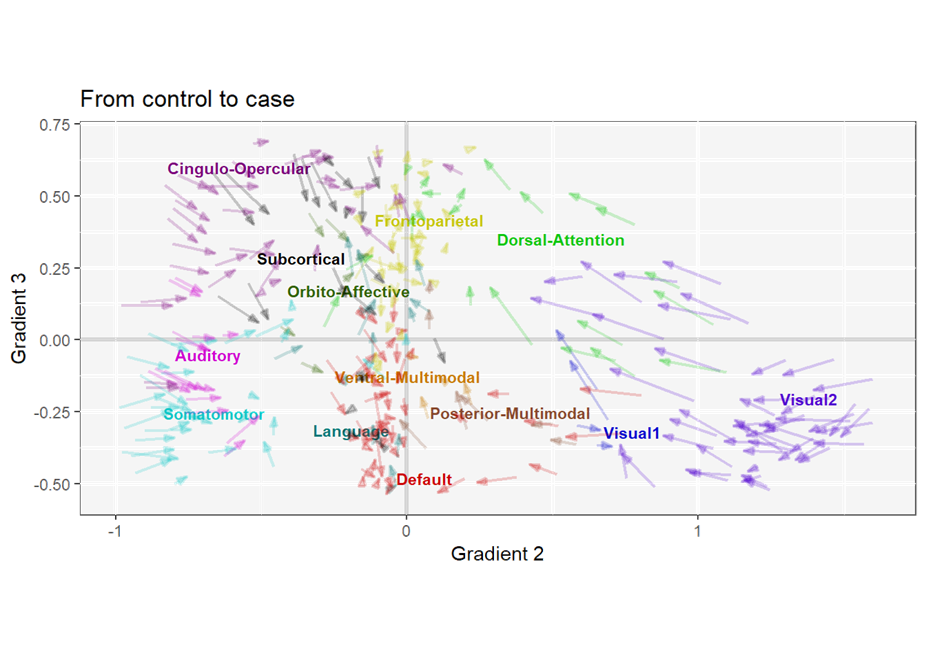
**Table S1.**



**Figure S2.** Boxplots of the eigenvalues from the first 10 dimensions extracted by the gradient analysis.



**Figure S3.** The scatter plot illustrates the group differences of each brain region on Gradients 1 and 2. In this plot, each arrow represents a region of interest (ROI) pointing from the mean gradient scores of the Control group to that of the SSD group. These arrows are colored by networks according to the Cole-Anticevic cortical atlas with the labels positioned at the mean gradient scores of each network. Specifically, this figure shows how ROIs from the default mode network (DMN), as compared to other networks, have decreased within-network, rather than between-network, segregation in SSD as all red arrows are pointing towards the network mean.

****

**Figure S4.** The scatter plot illustrates the group differences of each brain region on Gradients 2 and 3. In this plot, each arrow represents a region of interest (ROI) pointing from the mean gradient scores of the Control group to that of the SSD group. These arrows are colored by networks according to the Cole-Anticevic cortical atlas with the labels positioned at the mean gradient scores of each network. Specifically, this figure shows how ROIs from the default mode network (DMN), as compared to other networks, have decreased within-network, rather than between-network, segregation in SSD as all red arrows are pointing towards the network mean.

**References**

1. Chavez S, Viviano J, Zamyadi M, Kingsley PB, Kochunov P, Strother S, Voineskos A (2018): A novel DTI-QA tool: Automated metric extraction exploiting the sphericity of an agar filled phantom. *Magn Reson Imaging* 46: 28–39.

2. Hawco C, Viviano JD, Chavez S, Dickie EW, Calarco N, Kochunov P, *et al.* (2018): A longitudinal human phantom reliability study of multi-center T1-weighted, DTI, and resting state fMRI data. *Psychiatry Res Neuroimaging* 282: 134–142.

3. Kochunov P, Dickie EW, Viviano JD, Turner J, Kingsley PB, Jahanshad N, *et al.* (2018): Integration of routine QA data into mega‐analysis may improve quality and sensitivity of multisite diffusion tensor imaging studies. *Hum Brain Mapp* 39: 1015–1023.

4. Viviano JD, Buchanan RW, Calarco N, Gold JM, Foussias G, Bhagwat N, *et al.* (2018): Resting-State Connectivity Biomarkers of Cognitive Performance and Social Function in Individuals With Schizophrenia Spectrum Disorder and Healthy Control Subjects. *Biol Psychiatry* 84: 665–674.

5. Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, *et al.* (2019): fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat Methods* 16: 111–116.

6. Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS (2011): Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Front Neuroinformatics* 5. https://doi.org/10.3389/fninf.2011.00013

7. Zhang Y, Brady M, Smith S (2001): Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 20: 45–57.

8. Dale AM, Fischl B, Sereno MI (1999): Cortical Surface-Based Analysis. *NeuroImage* 9: 179–194.

9. Treiber JM, White NS, Steed TC, Bartsch H, Holland D, Farid N, *et al.* (2016): Characterization and Correction of Geometric Distortions in 814 Diffusion Weighted Images ((J. Najbauer, editor)). *PLOS ONE* 11: e0152472.

10. Jenkinson M, Bannister P, Brady M, Smith S (2002): Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage* 17: 825–841.

11. Dickie EW, Anticevic A, Smith DE, Coalson TS, Manogaran M, Calarco N, *et al.* (2019): Ciftify: A framework for surface-based analysis of legacy MR acquisitions. *NeuroImage* 197: 818–826.

12. Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, *et al.* (2013): The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* 80: 105–124.

13. Robinson EC, Jbabdi S, Glasser MF, Andersson J, Burgess GC, Harms MP, *et al.* (2014): MSM: A new flexible framework for Multimodal Surface Matching. *NeuroImage* 100: 414–426.

14. Krishnan A, Williams LJ, McIntosh AR, Abdi H (2011): Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. *Neuroimage* 56: 455–475.