Archival Report

Biological Psychiatry

Shared and Anxiety-Specific Pediatric Psychopathology Dimensions Manifest Distributed Neural Correlates

Julia O. Linke, Rany Abend, Katharina Kircanski, Michal Clayton, Caitlin Stavish, Brenda E. Benson, Melissa A. Brotman, Olivier Renaud, Stephen M. Smith, Thomas E. Nichols, Ellen Leibenluft, Anderson M. Winkler, and Daniel S. Pine

ABSTRACT

BACKGROUND: Imaging research has not yet delivered reliable psychiatric biomarkers. One challenge, particularly among youth, is high comorbidity. This challenge might be met through canonical correlation analysis designed to model mutual dependencies between symptom dimensions and neural measures. We mapped the multivariate associations that intrinsic functional connectivity manifests with pediatric symptoms of anxiety, irritability, and attention-deficit/hyperactivity disorder (ADHD) as common, impactful, co-occurring problems. We evaluate the replicability of such latent dimensions in an independent sample.

METHODS: We obtained ratings of anxiety, irritability, and ADHD, and 10 minutes of resting-state functional magnetic resonance imaging data, from two independent cohorts. Both cohorts (discovery: n = 182; replication: n = 326) included treatment-seeking youth with anxiety disorders, with disruptive mood dysregulation disorder, with ADHD, or without psychopathology. Functional connectivity was modeled as partial correlations among 216 brain areas. Using canonical correlation analysis and independent component analysis jointly we sought maximally correlated, maximally interpretable latent dimensions of brain connectivity and clinical symptoms.

RESULTS: We identified seven canonical variates in the discovery and five in the replication cohort. Of these canonical variates, three exhibited similarities across datasets: two variates consistently captured shared aspects of irritability, ADHD, and anxiety, while the third was specific to anxiety. Across cohorts, canonical variates did not relate to specific resting-state networks but comprised edges interconnecting established networks within and across both hemispheres.

CONCLUSIONS: Findings revealed two replicable types of clinical variates, one related to multiple symptom dimensions and a second relatively specific to anxiety. Both types involved a multitude of broadly distributed, weak brain connections as opposed to strong connections encompassing known resting-state networks.

https://doi.org/10.1016/j.biopsych.2020.10.018

The current study combines the dimensional assessment of psychiatric symptoms, potentially providing a better fit to neural measures than diagnostic categories (1–5) with resting-state functional magnetic resonance imaging (rsfMRI). We used advanced multivariate statistical techniques to identify highly correlated latent dimensions of psychopathology and brain connectivity. This is essential in identifying neural mechanisms that mediate clinical symptoms and thus represent appropriate targets for novel interventions. Such work is particularly needed among youth, in whom seeds of later-life psychopathology present as common, often co-occurring problems (6). Specifically, the study focuses on the neural correlates of pediatric irritability, attention-deficit/hyperactivity disorder (ADHD), and anxiety given prior evidence of both shared and distinct neural correlates among these symptom domains (1).

We use canonical correlation analysis (CCA) to simultaneously model dimensional clinical and neural measures (7,8). This approach might be more sensitive to complex relationships among symptom and neural data than alternative approaches. This includes approaches used previously that first model covariance structure among clinical dimensions before relating these latent symptom dimensions to preselected brain networks (1).

Recent studies applied CCA to clinical and rsfMRI data in adults and adolescents (9,10). The current study extends such work in three ways. First, while prior work applied CCA to rsfMRI data in treatment-seeking adults (10) and community-dwelling youth (9), we target treatment-seeking youth identified by clinicians. Second, prior work in youth confirmed that this method can differentiate well-established, but vastly distinct, clinical domains such as psychosis and emotional and

ISSN: 0006-3223

behavioral problems. Here, we focus on three more closely related and often comorbid domains: irritability, ADHD, and anxiety. We test a hypothesis consistent with previous work using other latent variable approaches combined with task-based fMRI (1,2): CCA yields latent phenotypes that capture both unique and shared aspects of irritability, ADHD, and anxiety. However, unlike past work, brain connectivity is not evoked by highly controlled tasks in the current study. Thus, more broadly distributed neural circuitry correlates are expected in the current study, as compared with correlates in previous studies.

Finally, as a third extension of past work, we evaluate the latent variables' replicability using novel sampling and analytic techniques. Prior CCA studies find replicable associations when discovery and replication cohorts represent subsets of the same sample (9,10) but not when they arise from independent cohorts (11). Robustness against sampling variability is essential for clinical applications of CCA, which possesses exploratory components that can make replication difficult. Thus, the current study used data from 2 independent cohorts of treatment-seeking youth assessed with similar methods. We treated the smaller sample (n = 182) as the discovery dataset, because it was assessed with homogeneous imaging parameters. The larger cohort (n = 326), assessed with heterogeneous imaging parameters, served as a replication dataset (12). Moreover, we employed analytic techniques that leverage independent component analysis (ICA) to improve interpretability of the canonical variates (13). Finally, we used a novel, stepwise permutation scheme (14) that addresses limitations in other CCA studies concerning the handling of nuisance variables and possible inflation of type I errors.

METHODS AND MATERIALS

Participants

Both samples comprised healthy volunteers and youth diagnosed with an anxiety disorder, disruptive mood dysregulation disorder, or ADHD by licensed clinicians using the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS) (15). Exclusion criteria were neurological disorders, autism and bipolar spectrum disorders, psychosis, substance use, MRI contraindications, and Full Scale IQ $<$ 70. Anxiety was assessed by using the parent- and youth-reported ratings of the five subscales of the Screen for Child Anxiety Related Disorders (16). Irritability was assessed with the first six items of the parent- and youth-reported Affective Reactivity Index (17). Parents quantified ADHD symptoms such as inattention and disruptive behavior through seven items assessed with the ADHD subscale of the Conners (18) in the discovery sample and the Child Behavior Checklist (19) in the replication sample. These 29 ratings of anxiety, irritability, and disruptive behavior (18 parent-reported, 11 self-reported) were used as input for the joint CCA+ICA.

Samples were similar in terms of sex ratios, proportions of anxiety disorders, oppositional defiant disorder, medication-free-to-medication-use ratios, and levels of parent-reported symptoms of irritability, ADHD, and anxiety. However, the discovery sample was older and had a higher IQ, a lower proportion of ADHD cases, a higher proportion of diagnosis-free and disruptive mood dysregulation disorder cases, and lower self-reported irritability and anxiety. Both cohorts were ethnically diverse and were recruited from urban, semi-rural, and rural areas (Figure 1, Supplement).

Acquisition and Preprocessing of Imaging Data

Discovery sample data were acquired at one site with two identical 3T General Electric Signa scanners (GE Healthcare, Chicago, IL). The replication sample comprised data from a 1.5T Siemens Avanto scanner (Siemens Healthineers, Erlangen, Germany), a 3T Siemens Tim Trio scanner, and a 3T Siemens Prisma scanner. A high-resolution T1-weighted structural image and 10 minutes of blood oxygen level–dependent changes during rest were collected from all participants, although sequences varied between samples and within the replication dataset (Supplement). Quality of the imaging data was assessed using MRIQC (20). The automated pipeline FMRIPREP (21) was used for preprocessing. We refrained from motion scrubbing and used instead ICA-AROMA, which reduces motion-related artifacts at least as well (22,23).

Functional Connectivity

The rsfMRI-connectivity network comprised 216 nodes derived from a 200-region parcellation scheme (24), augmented by eight subcortical regions per hemisphere obtained using FreeSurfer segmentation (nucleus accumbens, nucleus caudatus, pallidum, putamen, amygdala, hippocampus, thalamus, and ventral diencephalon) (25). Framewise displacement and spatial standard deviation of the temporal difference data (26), but not global signal, were regressed out from the time series. Functional connectivity was quantified using partial correlations, which offer an estimate of direct (as opposed to indirect or shared) connectivity between each pair of nodes (edges). Because the resulting network matrices are symmetric, only half of the edges (i.e., 23,220) were analyzed.

CCA, ICA, and Permutation Testing

Covariates (age, sex, race, IQ, psychotropic medication, and scanner for the discovery sample; additionally, site and sequence type for the replication sample) were regressed out from both imaging and clinical variables before dimensionality reduction. All 29 symptom ratings were included; dimensionality of rsfMRI was reduced using principal component analysis (PCA) before CCA (7). Residuals were projected to a lower dimensional space where data are exchangeable, thus mitigating spurious dependencies among observations introduced by residualization (14).

Given two sets of variables (here, imaging data [Y] and symptom ratings [X]), CCA seeks linear mixtures within each set (i.e., canonical variables [CVs]; $U = Y \times A$ and $V = X \times B$), such that each resulting mixture (U) from one set is maximally correlated with a corresponding mixture (V) from the other set, but uncorrelated with all other mixtures in either set. We use upper-case letters U and V to represent the whole set of canonical variables on the imaging and clinical side, respectively, and lower-case letters followed by subscripts to indicate the order of the canonical correlations, from higher to smaller, u_k and v_k, to denote specific latent variables.

Small perturbations in the original data could possibly lead to arbitrary rotations of the CCA solutions. To mitigate the

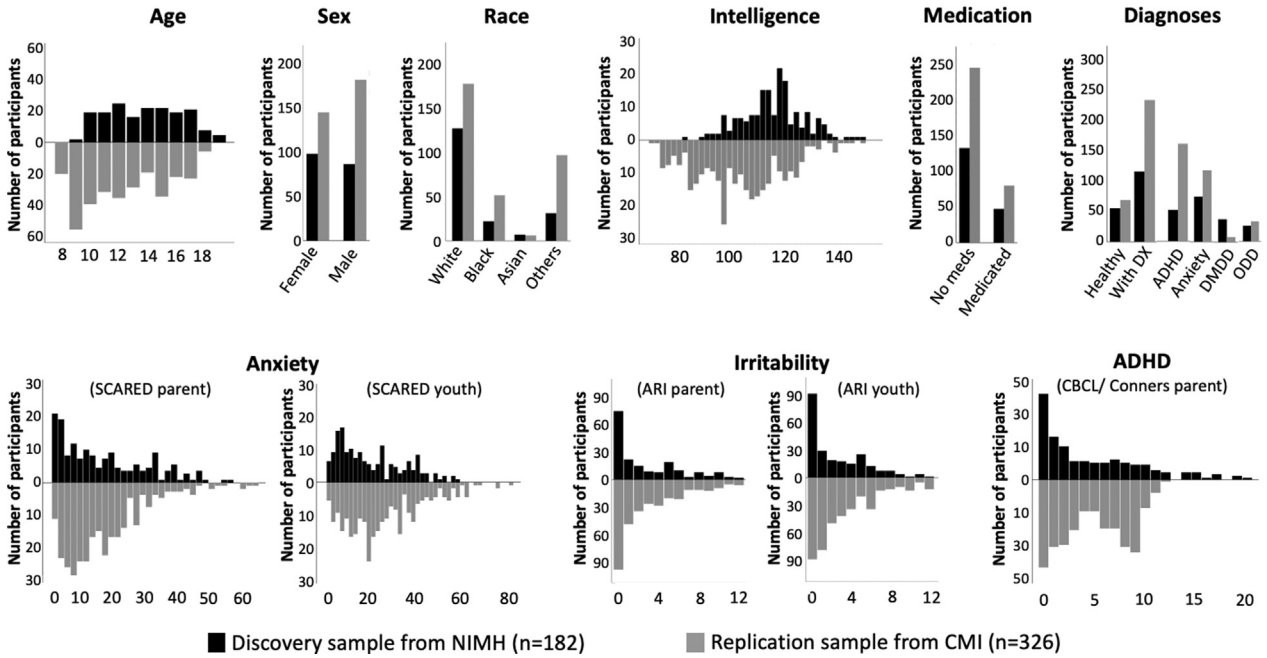


Figure 1. Demographics and clinical characteristics of the discovery and replication samples. ADHD, attention-deficit/hyperactivity disorder; ARI, Affective Reactivity Index; CBCL, Child Behavior Checklist; CMI, Child Mind Institute; DMDD, disruptive mood dysregulation disorder; DX, diagnosis; NIMH, National Institute of Mental Health; ODD, oppositional defiant disorder; SCARED, Screen for Child Anxiety Related Disorders.

problem and aid interpretability, we subjected the stacked CVs to ICA, seeking CVs that were not only orthogonal but also statistically independent. The joint CCA+ICA procedure was performed using a modification of a recently proposed algorithm for permutation inference for CCA (14), thus allowing not only characterization and better disambiguation of the resulting CVs but also valid statistical inference (details in the Supplement). Below, where we refer to results of CCA, these are to be understood as results of the joint inference using CCA+ICA.

Because the number of CVs is determined by the smallest input dataset, we obtained 29 CVs. Statistical significance was determined using 10,000 permutations. In the permutation test, for each estimated CV (post-ICA), variance already explained by CVs with stronger, significant canonical correlations were removed in an iterative procedure (14). Canonical correlations were considered significant at alpha = .05 after familywise error rate (FWER) correction using a closed testing procedure. The nonsymmetric redundancy index (27), which gives the mean variance of the clinical data explained by imaging data, and vice versa, is reported in the Supplement.

Replicability

Replicability of the CVs was determined based on three criteria: 1) stability within the same dataset across variations in the number of PCA components that entered CCA relative to the sample size (input-to-participant ratio), 2) similarities of latent clinical patterns, and 3) similarities of latent connectivity patterns identified independently in the two samples. A prior CCA study in youth reported replicability only for clinical but not for rsfMRI patterns (9); thus, we decided to evaluate the replicability of the clinical and the connectivity patterns as separate criteria.

To evaluate the first criterion, we performed three analyses that varied the input-to-participant ratios. The primary analyses used an input-to-participant ratio of 1:2, which translated into 64 rsfMRI components, explaining 75% of the between-subject variance in rsfMRI connectivity in the discovery sample. In the replication cohort, 134 rsfMRI components were used; these explained only 57% of the variance, possibly owing to more unstructured noise in this dataset. This primary analysis was supplemented by two secondary analyses using input-to-participant ratios of 1:3 and 1:4, thereby reducing risks of overfitting, at the expense of explaining less variance. This was accomplished by using fewer imaging principal components as input to the CCA. Results were compared across the three ratios by examining cross-correlations among CCA components (e.g., corr[$v_{1|1:2}$, $Y_D \times a_{1|1:3}$] and corr[$u_{1|1:2}$, $X_D \times b_{1|1:3}$]). Statistical significance was determined using 10,000 permutations, with a threshold of $p_{FWER} < .05$ within each set of comparisons. Because psychiatric symptoms might relate to components that explain relatively little variance in the imaging data, we also discuss CVs that solely replicated at the 1:3 ratio but could be found in the replication cohort.

To test the second and third criteria, we used joint CCA+ICA in the replication dataset. Canonical weights from each dataset were applied to the input data from the other dataset; these products were then correlated with the CVs identified in that dataset (e.g., corr[$v_{1|D}$, $Y_D \times a_{1|R}$] and corr[$u_{1|D}$, $X_D \times b_{1|R}$]). Clinical and connectivity patterns were considered replicable when both the application of weights from the discovery to the replication dataset and the application of weights from the replication to discovery dataset yielded statistically significant

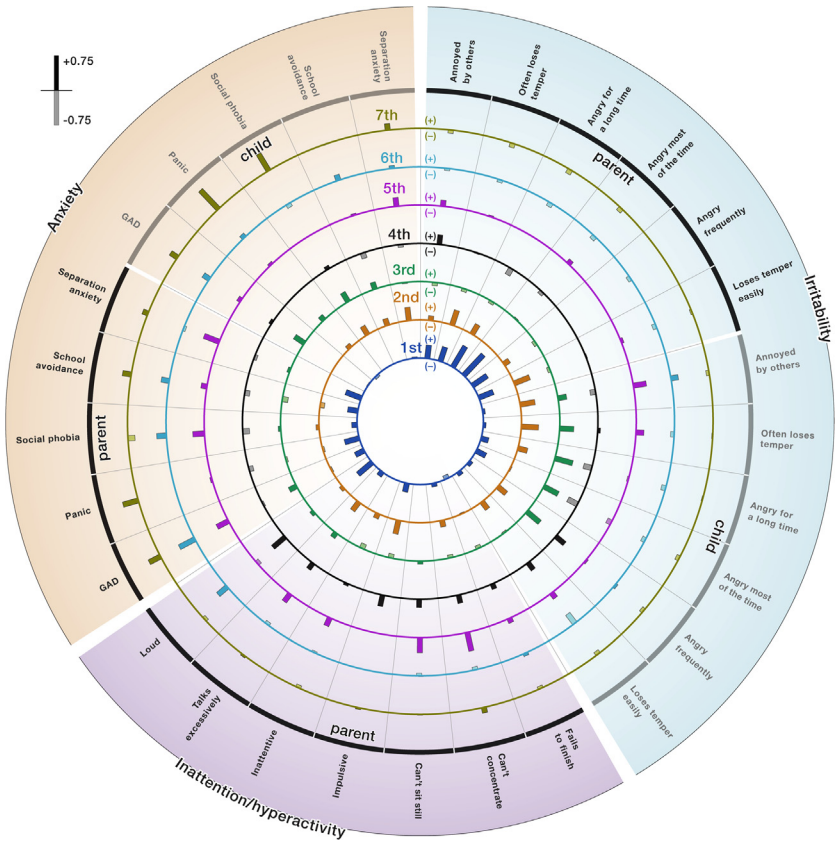


Figure 2. Clinical loadings in the discovery dataset. Associations between the latent dimensions and symptoms in the domains of anxiety, irritability, and behavioral problems. Each of the seven concentric circles shows the positive (solid fill) and negative correlations (transparent fill) between the canonical variate and the original symptom ratings as bars. The length of the bar indicates the strength of the association. Exact numbers of the loadings are provided in Table S3. Canonical correlation analysis results are characterized by sign indeterminacy, meaning that it is valid to flip the sign for an entire latent dimension, which will affect the directions of the correlations.

associations. We used 10,000 permutations to establish significance. However, thresholds differed for clinical and connectivity patterns. We used a stringent threshold of $p_{FWER} < .05$ to determine replicability of the imaging and clinical patterns; additionally, we also investigated a more lenient threshold of $p_{uncorr} < .05$ for replicability of the connectivity pattern. This decision was motivated by two factors. First, a prior CCA study finding replicable clinical patterns did not report replicable connectivity patterns across two subsets of a single sample (9). This raises questions as to whether any evidence of replicability can be detected with even liberal statistical thresholds. Second, in the current study, significant differences exist between cohorts in all metrics, quantifying the quality of the imaging data (Figure S1); this contrasts with the broadly similar profiles for symptom ratings (Table S1 and Figure S1).

Interpretation of CVs

To interpret the significant CVs, we investigated their correlations with the residualized input data. These correlations between latent variables derived from the CCA and input data (i.e., symptom ratings, connectivity matrices) are henceforth referred to as canonical loadings. Consistent with prior studies (7,28), we focused on clinical items with loadings $|r| > 0.2$, resembling a small to moderate effect, to interpret and label key CVs. However, we extended this approach by limiting our focus to replicating clinical loadings, i.e., loadings $|r| > 0.2$ that could be observed across samples. Similarly, we emphasized edge loadings that replicated across samples. However, given the differences in the quality of the imaging data across samples, we applied a more lenient threshold of $|r| > 0.15$ to the replication cohort.

RESULTS

In the discovery cohort, seven CVs associated symptoms with rsfMRI-connectivity ($CV_{1|D}$: $r = .74$, $p_{FWER} = .0029$; $CV_{2|D}$: $r = .73$, $p_{FWER} = .0045$; $CV_{3|D}$: $r = .73$, $p_{FWER} = .0092$; $CV_{4|D}$: $r = .70$, $p_{FWER} = .0175$; $CV_{5|D}$: $r = .69$, $p_{FWER} = .0245$; $CV_{6|D}$: $r = .69$, $p_{FWER} = .0340$; $CV_{7|D}$: $r = .68$, $p_{FWER} = .0497$) (Figure S2). These seven CVs explained 27.7% of symptom-level variance and 8.1% of rsfMRI-connectivity variance. Joint CCA+ICA in the replication dataset generated five CVs ($CV_{1|R}$: $r = .75$, $p_{FWER} = .0198$; $CV_{2|R}$: $r = .72$, $p_{FWER} = .0311$; $CV_{3|R}$: $r = .72$, $p_{FWER} = .0321$; $CV_{4|R}$: $r = .72$, $p_{FWER} = .0364$; $CV_{5|R}$: $r = .71$, $p_{FWER} = .0440$) (Figure S11) that represented 18.8% of the variance in the clinical data and 0.9% of the variance in the rsfMRI-connectivity data.

Our key hypothesis concerned identifying both cross-dimension and specific variates. Consistent with this hypothesis, latent clinical phenotypes ($v_{1\text{-}7|D}$ and $v_{1\text{-}5|R}$) could be differentiated in terms of specificity of associated symptoms. In both datasets, we observed latent variables that loaded

$|r| > 0.2$ exclusively on several items from the anxiety domain ($v_{7|D}$, $v_{1|R}$) and others capturing aspects shared across two ($v_{3|D}$, $v_{4|D}$, $v_{3|R}$) or all three domains ($v_{1|D}$, $v_{2|D}$, $v_{5|D}$, $v_{6|D}$, $v_{2|R}$, $v_{4|R}$, $v_{5|R}$) (Figure 2, Figure S12, Tables S7 and S12).

All latent variables ($u_{1\text{-}7|D}$ and $u_{1\text{-}5|R}$) involved distinct, albeit broadly distributed, connectivity patterns with many connections between well-known resting-state networks within and across both hemispheres. All $u_{1\text{-}7|D}$ and $u_{1\text{-}5|R}$ showed equal numbers of negative and positive correlations with edges. Connectivity patterns were denser in the discovery relative to the replication dataset applying a threshold of $|r| > 0.2$ (discovery dataset: $u_{1|D}$: 1404 connections, $u_{2|D}$: 2068 connections, $u_{3|D}$: 5602 connections, $u_{4|D}$: 2490 connections, $u_{5|D}$: 2968 connections, $u_{6|D}$: 2656 connections, and $u_{7|D}$: 3226 connections; replication dataset: $u_{1|R}$: 150 connections; $u_{2|R}$: 130 connections, $u_{3|R}$: 74 connections; $u_{4|R}$: 74 connections; $u_{5|R}$: 132 connections) (Figures 3–5, Figures S5–S10, S13–S17).

We highlight three CVs passing all three replicability criteria highlighted in the Methods and Materials section and two latent dimensions passing only the first two replicability criteria (stability within the same dataset and replicability of clinical patterns). Full results concerning replicability appear in the Supplement (Tables S8–S11, S13–S20). We describe replicable CVs based on the specificity of the clinical patterns ranging from shared between all three clinical domains to anxiety-specific.

Irritability, Anxiety, and ADHD ($CV_{2|D}$, $CV_{3|D}$, and $CV_{5|R}$)

In both samples, $CV_{2|D}$, $CV_{3|D}$, and $CV_{5|R}$ were robust to variations in the participant-to-input ratios (all $r > .67$, all $p_{FWER} = .0001$) (Tables S8–S11, S13–S16). Clinical patterns associated with $v_{3|D}$ and $v_{5|R}$ were negatively associated across samples (corr[$v_{5|R}$, $X_R \times b_{3|D}$]: $r = -.20$, $p_{uncorr} = .0006$, $p_{FWER} = .0175$; corr[$v_{3|D}$, $X_D \times b_{5|R}$]: $r = -.19$, $p_{uncorr} = .0141$, $p_{FWER} = .3797$). Both $v_{3|D}$ and $v_{5|R}$ loaded >.20 on the irritability and anxiety domains, where close inspection suggested informant effects; $v_{3|D}$ captured youth-reported whereas $v_{5|R}$ loaded on parent-reported irritability and anxiety (Figure 3, Tables S7 and S12). Yet, connectivity patterns correlated across samples (corr[$u_{5|R}$, $Y_R \times a_{3|D}$]: $r = .12$, $p_{uncorr} = .0375$, $p_{FWER} = .7425$; corr[$u_{3|D}$, $Y_D \times a_{5|R}$]: $r = .32$, $p_{uncorr} = .0001$, $p_{FWER} = .0007$). Inspection of the connectivity loadings showed that in both samples this transdimensional phenotype was associated with edges interconnecting established resting-state networks within and across both hemispheres (Figure S18).

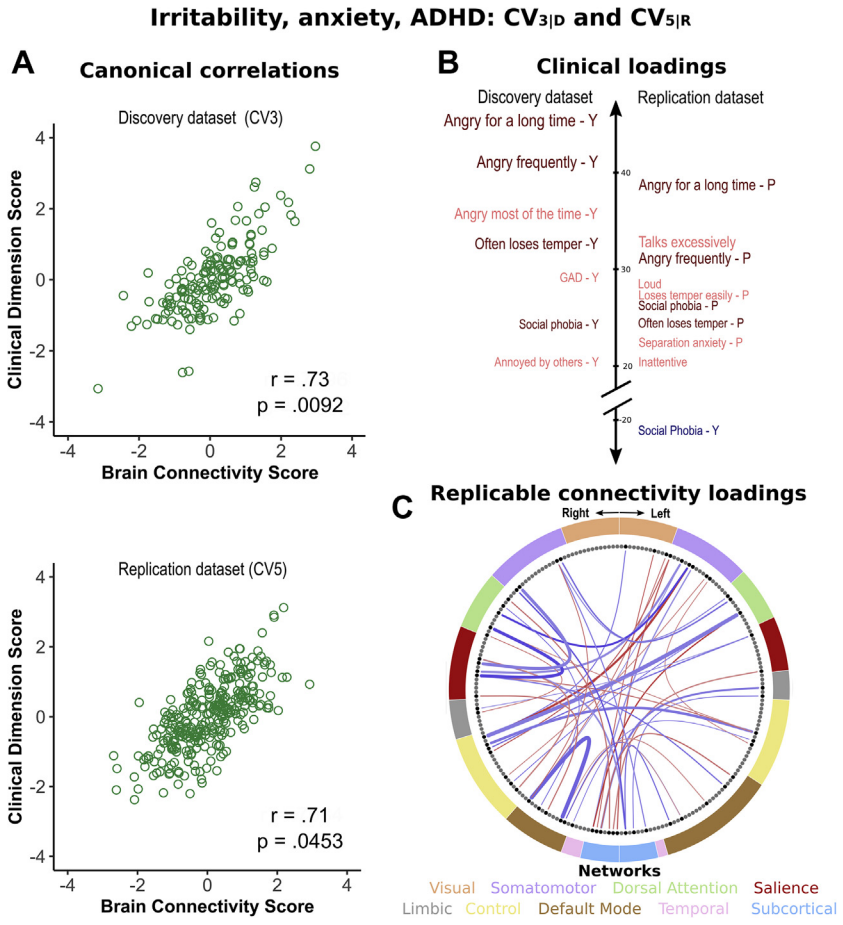


Figure 3. Replicable, transdimensional latent variable ($CV_{3|D}$, $CV_{5|R}$). **(A)** Scatter plots show canonical variate (CV) 3 from the discovery dataset (D) and 5 from the replication dataset (R), which represent linear combinations of brain connectivity scores obtained during resting-state functional magnetic resonance imaging in the horizontal axis, and linear combinations of clinical scores derived from symptom ratings in the vertical axis. **(B)** Clinical loadings $|r| > 0.2$ in for both datasets, showing the same symptoms but an informant effect. Dark red indicates symptoms associated with the latent dimension in both datasets. **(C)** Edges in red that load strongly positively on $u_{3|D}$ and $u_{5|R}$. Edges that load strongly negatively on $u_{3|D}$ and $u_{5|R}$ are depicted in blue. Given baseline differences in the strength of the connectivity patterns, connectivity maps were thresholded at $|r| > 0.2$ for the discovery sample and at $|r| > 0.15$ for the replication sample. Only edges that loaded highly positively or negatively in both datasets were retained for this figure. ADHD, attention-deficit/hyperactivity disorder; P, parent; Y, youth.

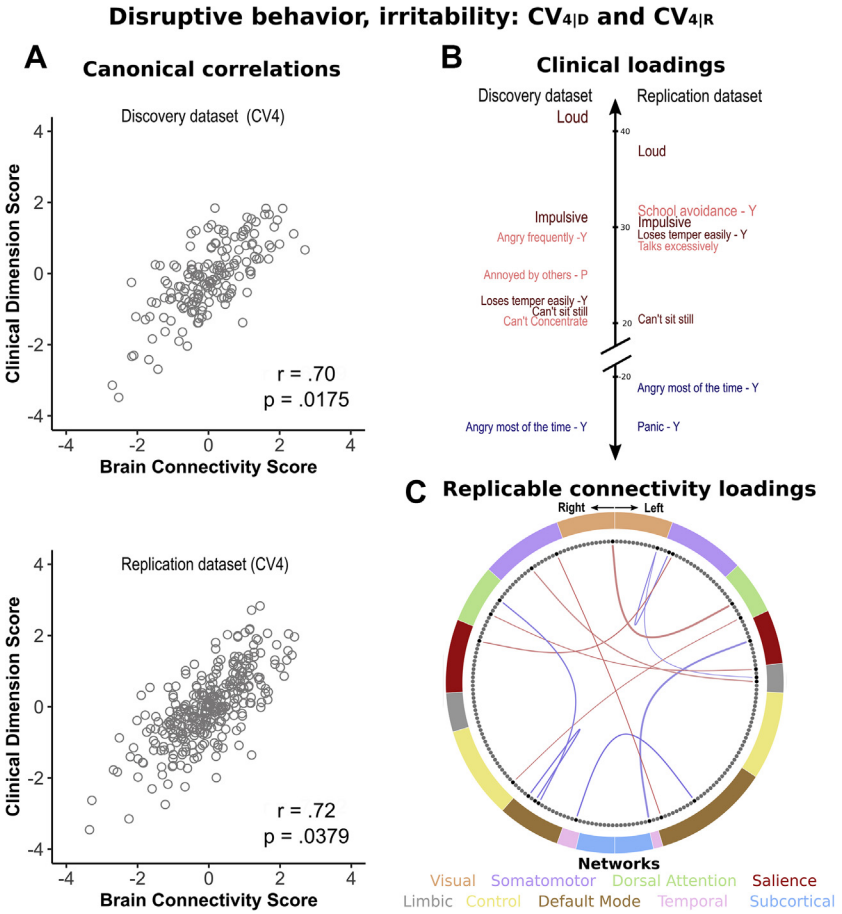


Figure 4. Replicable, shared aspects of disruptive behavior and irritability ($CV_{4|D}$, $CV_{4|R}$). **(A)** Scatter plots show canonical variate (CV) 4 from the discovery dataset (D) and 4 from the replication dataset (R), which represent linear combinations of brain connectivity scores obtained during resting-state functional magnetic resonance imaging in the horizontal axis, and linear combinations of clinical scores derived from symptom ratings in the vertical axis. **(B)** Clinical loadings $|r| > 0.2$ in for both datasets. Dark red indicates symptoms associated with the latent dimension in both datasets. **(C)** Edges in red that load strongly positively on $u_{4|D}$ and $u_{4|R}$. Edges that load strongly negatively on $u_{4|D}$ and $u_{4|R}$ are depicted in blue. Given baseline differences in the strength of the connectivity patterns, connectivity maps were thresholded at $|r| > 0.2$ for the discovery sample and at $|r| > 0.15$ for the replication sample. Only edges that loaded highly positively or negatively in both datasets were retained for this figure. P, parent; Y, youth.

Interestingly, $v_{2|D}$ was also robust against variations in participant-to-input ratios and showed substantial positive loadings >.20 on the same three parent-report items from the irritability domain ("Often loses temper," "Angry for a long time," "Loses temper easily") and one from the ADHD domain as $v_{5|R}$ ("Talks excessively") (Figure 2 and Figure S12, Tables S7 and S12). Moreover, clinical loadings for $CV_{2|D}$ and $CV_{5|R}$ significantly correlated across cohorts (corr[$v_{5|R}$, $X_R \times b_{2|D}$]: $r = .25$, $p_{uncorr} = .0001$, $p_{FWER} = .0004$; corr[$v_{2|D}$, $X_D \times b_{5|R}$]: $r = .45$, $p_{uncorr} = .0001$, $p_{FWER} = .0001$). However, within each sample, connectivity patterns associated with the two latent phenotypes were different, although brain connectivity data informed latent clinical dimensions.

Disruptive Behavior and Irritability ($CV_{4|D}$ and $CV_{4|R}$)

In both cohorts, $CV_{4|D}$ and $CV_{4|R}$ were robust to input-to-participant ratio variations (all $r > .31$, $p_{FWER} = .0001$) (Tables S8–S11, S13–S16). Similarities between $CV_{4|D}$ and $CV_{4|R}$ arose when applying clinical weights from the discovery to the replication dataset (corr[$v_{4|R}$, $X_R \times b_{4|D}$]: $r = .19$, $p_{uncorr} = .0006$, $p_{FWER} = .0229$) and vice versa (corr[$v_{4|D}$, $X_D \times b_{4|R}$]: $r = .27$, $p_{uncorr} = .0009$, $p_{FWER} = .0190$). Connectivity patterns were also associated using an uncorrected threshold, when applying weights from the replication to the discovery sample (corr[$u_{4|D}$, $Y_D \times a_{4|R}$]: $r = .19$, $p_{uncorr} = .0108$, $p_{FWER} = .3494$).

Across samples, $CV_{4|D}$ and $CV_{4|R}$ loaded >.20 on three items characterizing disruptive behavior from the ADHD domain ("Can't sit still," "Impulsive," "Loud") and one item from the domain of irritability ("Loses temper easily"). Furthermore, both $CV_{4|D}$ and $CV_{4|R}$ loaded negatively on one irritability item ("Angry most of the time") (Figure 4, Tables S7 and S12). Inspection of substantial edge loadings in both samples indicated strong representations in the variate of connections among nodes in motor, attention, default mode, and temporal-parietal networks (Figure S19).

Anxiety ($CV_{7|D}$ and $CV_{1|R}$, $CV_{3|R}$)

The last set of replicable CVs comprised $CV_{7|D}$ in the discovery cohort, which correlated with both $CV_{1|R}$ and $CV_{3|R}$ in the replication data set. All three CVs emerged in analyses using input-to-participant ratios of 1:2 and 1:3 (all $r > .31$, all $p_{FWER} = .0001$) (Tables S8–S11, S13–S16). Associations manifested between $v_{7|D}$ and $v_{3|R}$ for clinical (corr[$v_{7|D}$, $X_D \times b_{3|R}$]: $r = .31$, $p_{uncorr} = .0001$, $p_{FWER} = .0017$; corr[$v_{3|R}$, $X_R \times b_{7|D}$]: $r = .44$, $p_{uncorr} = .0001$, $p_{FWER} = .0001$) and for connectivity patterns, when applying an uncorrected threshold (corr[$u_{7|D}$, $Y_D \times a_{3|R}$]: $r = .19$, $p_{uncorr} = .0136$, $p_{FWER} = .4120$; corr[$u_{3|R}$, $Y_R \times a_{7|D}$]: $r =$

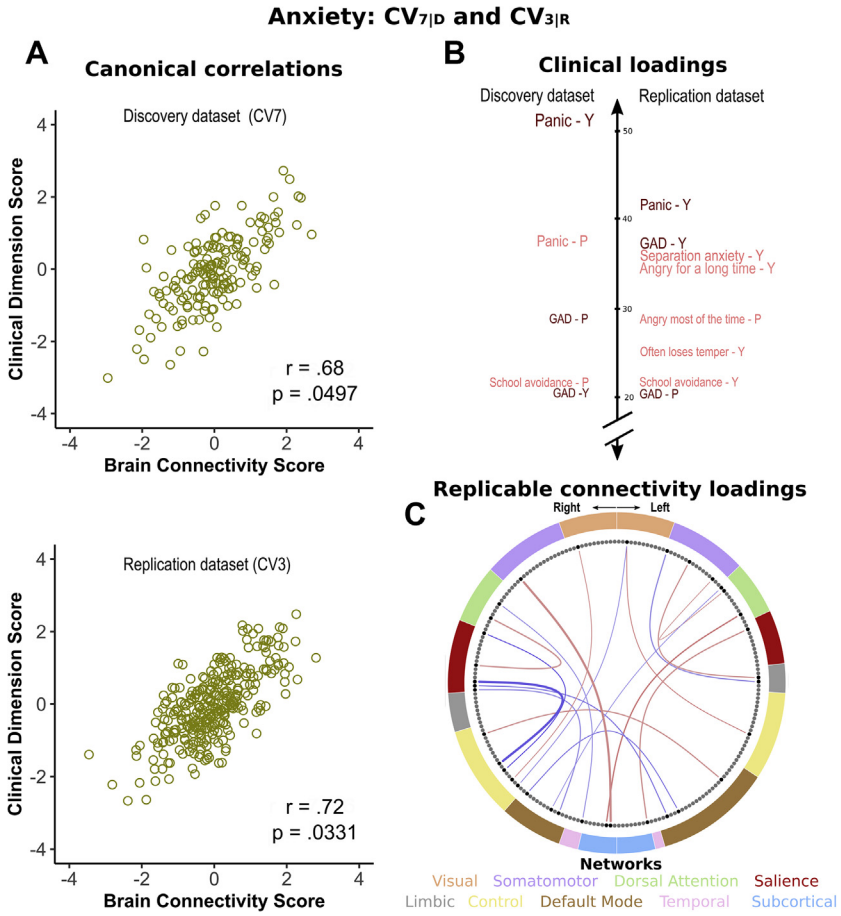


Figure 5. Replicable, anxiety-specific latent variable ($CV_{7|D}$, $CV_{3|R}$). **(A)** Scatter plots show canonical variate (CV) 7 from the discovery dataset (D) and 3 from the replication dataset (R), which represent linear combinations of brain connectivity scores obtained during resting-state functional magnetic resonance imaging in the horizontal axis, and linear combinations of clinical scores derived from symptom ratings in the vertical axis. **(B)** Clinical loadings $|r| > 0.2$ in for both datasets. Dark red indicates symptoms associated with the latent dimension in both datasets. **(C)** Edges in red load strongly positively on $u_{7|D}$ and $u_{3|R}$. Edges that load strongly negatively on $u_{7|D}$ and $u_{3|R}$ are depicted in blue. Given baseline differences in the strength of the connectivity patterns, connectivity maps were thresholded at $|r| > 0.2$ for the discovery sample and at $|r| > 0.15$ for the replication sample. Only edges that loaded highly positively or negatively in both datasets were retained for this figure. GAD, generalized anxiety disorder; P, parent; Y, youth.

.10, p_{uncorr} = .0696, p_{FWER} = 1). For $V_{7|D}$ and $v_{3|R}$, both variates loaded >.20 on the same three anxiety items ("parent-reported generalized anxiety disorder," "youth-reported generalized anxiety disorder," "youth-reported panic") (Figure 5, Tables S7 and S12). Replicable edges connected subcortical structures with the dorsal-attention and motor network as well as the control and default mode networks with sensory, motor, and attention networks (Figure S20).

Similarities in the clinical patterns were also observed between $v_{7|D}$ and $v_{1|R}$. Both loaded >.20 on four items measuring anxiety ("parent-reported generalized anxiety disorder," "parent-reported panic," "parent-reported school avoidance," "youth-reported panic"; corr[$v_{7|D}$, $X_D \times b_{1|R}$]: r = .45, p_{uncorr} = .0001, p_{FWER} = .0001; corr[$v_{1|R}$, $X_R \times b_{7|D}$]: r = .51, p_{uncorr} = .0001, p_{FWER} = .0001) (Figure 2 and Figure S12; Tables S7 and S12). However, unlike for $u_{3|R}$, associated connectivity patterns between $u_{7|D}$ and $u_{1|R}$ were uncorrelated even when the uncorrected threshold was applied.

DISCUSSION

Three key findings emerge from this study. First, analyses found seven CVs in a discovery dataset; four showed stability within the discovery dataset and replicability of clinical patterns in an independent sample; three CVs demonstrated at least weak signs of replicability for the associated rsfMRI connectivity patterns. This suggests the presence of meaningful relations between patterns of intrinsic brain connectivity and psychiatric symptom dimensions in youth. Second, the three most strongly replicable CVs from the discovery dataset varied in clinical specificity; one loaded on all three domains, the second captured shared aspects of irritability and ADHD, and the third loaded specifically on anxiety. Finally, CVs showed weak to modest associations, with multiple edges spanning widely distributed brain areas.

Pediatric psychopathology involves broadly correlated symptom dimensions (1–6). Dimensions of irritability, ADHD, and anxiety are particularly closely interrelated. Understanding of these cross-dimension relations may follow from research on shared and unique neural correlates. Past work in this area assessed symptom covariation independent of imaging data before then relating symptoms to task-based imaging patterns (1). CCA connects clinical and neural measures simultaneously to identify more complex relations (7,8). We used rating scales employed in the previous task-based fMRI research examining unique and shared dimensions of pediatric psychopathology (1). Using these measures, the current rsfMRI study identified two variates loading strongly on

multiple clinical dimensions and a third loading strongly only on anxiety items. Thus, consistent with our hypotheses based on past studies, current findings demonstrate coexisting cross-dimensional and domain-specific neural correlates in treatment-seeking youth.

The detection of only anxiety but not irritability or ADHD-specific neural correlates in the current study could reflect many factors. These include differences between task-based and rsfMRI methods, differential sensitivity in CCA to particular domain-specific features, or biological features of anxiety that generate specific rsfMRI signatures. Additional imaging research might seek to refine clinical groupings based on replicable cross-study patterns for these and other interrelated dimensions.

Findings in the current and past CCA studies exhibited both similarities and differences. Cross-sample correlations for clinical loadings in the current study were notably similar in magnitude to those for variables involving emotion symptoms in the only other study of cross-domain pediatric psychopathology (9). Given differences across the two studies, such consistency speaks to the robust nature of pediatric emotional problem manifestations. The previous study also found strong cross-sample replicability for a pure externalizing factor, which did not emerge in the current study. Failure to detect this factor might reflect lesser diversity in targeted symptoms or larger proportions of treatment-seeking cases in the current study. Finally, unlike past research in treatment-seeking adults, the current study showed cross-sample replicability of latent clinical and connectivity patterns, a finding that might reflect age-related differences or distinct analytic approaches.

Interesting rsfMRI patterns manifested. Connectivity related to clinical dimensions was broadly distributed, involving hundreds of relatively weakly loading interhemispheric and within-hemisphere connections spanning distinct networks. Moreover, while within-sample stability was acceptable in the discovery sample, rsfMRI patterns minimally correlated across datasets. Interestingly, such weak replicability manifested alongside stronger replicability for clinical patterns, themselves defined by relations with rsfMRI. Replicable clinical patterns defined by less replicable rsfMRI patterns raise important questions for future studies. First, greater cross-sample differences existed for the fMRI than clinical assessments. Thus, whether homogeneous cross-sample imaging methods could generate improved rsfMRI replicability remains unclear. Second, replicable clinical patterns defined by minimally replicable fMRI patterns could arise from "many-to-one" mappings between neural and clinical variables. Such configurations commonly underlie brain-behavior relationships at many spatial scales. Thus, whether such "many-to-one" patterns also represent a common motif for mental disorders remains unclear.

From the clinical perspective, broadly distributed connectivity disturbances might require a diverse set of approaches to identify targets for novel interventions. Currently, therapies such as cognitive training or neural stimulation target functions in specific networks (29–31). However, at least for pediatric anxiety, irritability, and disruptive behavior, broadly distributed patterns may better represent the nature of connectivity disturbances during rest than patterns limited to particular networks. The focus on broad connectivity disturbances as opposed to particular networks might increase effect sizes of studies relating clinical domains to intrinsic brain connectivity.

Findings inform analytic decisions in future CCA studies. Different analyses within and across samples used different rsfMRI data, accounted for different amounts of overall rsfMRI variance, and yielded differences in CV structure. That input affects output is not unique to CCA. However, no ground truth informs selection of PCA-based or other input components for CCA. Thus, risk of overfitting is balanced against risk of omitting relevant variance through dimensionality reduction. Overfitting is reduced by ensuring proportionally more research participants than variables (32,33). However, particularly in moderately sized datasets, dimensionality reduction can exclude rsfMRI variance components that, even if small, powerfully relate clinical dimensions to connectivity patterns. Such factors create challenges that likely impact findings. The presence of modestly replicable clinical loadings across analyses in the current study suggests the promise of continued iterative work targeting these challenges.

One major limitation of the current study are the medium sample sizes. Also, differences in scanners, imaging-acquisition parameters, and data-quality indices introduced noise that decreased the probability of fully replicating findings across datasets. In effect, larger proportions of rsfMRI connectivity variance are explained in the smaller but homogeneous discovery sample, as evidenced by PCA. Furthermore, we did not include youth ratings of ADHD symptoms or ratings of depressive symptoms, another highly prevalent symptom dimension in youth.

Our findings implicate co-occurring transdimensional and anxiety-specific neural features in pediatric psychopathology. Results further suggest that pediatric clinical dimensions reflect widely distributed brain connectivity patterns. Thus, as with genetic correlates, neural correlates of some pediatric psychopathology dimensions may reflect hundreds of individually small associations.

ACKNOWLEDGMENTS AND DISCLOSURES

This work is supported by the National Institute of Mental Health (NIMH) Intramural Research Program (Grant Nos. ZIAMH002786 [to EL], ZIAMH002778 [to EL], and ZIAMH002782 [to DSP]), conducted under NIH Clinical Study Protocols described at ClinicalTrials.gov (NCT02531893, NCT00025935, and NCT00018057).

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Emotion and Development Branch (JOL, RA, KK, MC, CS, BEB, MAB, EL, AMW, DSP), National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland; Methodology and Data Analysis (OR), Department of Psychology, University of Geneva, Switzerland; and the Wellcome Trust Centre for Integrative Neuroimaging (SMS) and the Big Data Institute (TEN), University of Oxford, Oxford, United Kingdom.

AMW and DSP contributed equally to this work.

Address correspondence to Julia O. Linke, Ph.D., at julia.linke@nih.gov.

Received Mar 13, 2020; revised Oct 16, 2020; accepted Oct 27, 2020.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.biopsych.2020.10.018.

REFERENCES

1. Kircanski K, White LK, Tseng W-L, Wiggins JL, Frank HR, Sequeira S, *et al.* (2018): A latent variable approach to differentiating neural mechanisms of irritability and anxiety in youth. JAMA Psychiatry 75:631–639.
2. Shanmugan S, Wolf DH, Calkins ME, Moore TM, Ruparel K, Hopson RD, *et al.* (2016): Common and dissociable mechanisms of executive system dysfunction across psychiatric disorders in youth. Am J Psychiatry 173:517–526.
3. Kaczkurkin AN, Moore TM, Calkins ME, Ciric R, Detre JA, Elliott MA, *et al.* (2018): Common and dissociable regional cerebral blood flow differences associate with dimensions of psychopathology across categorical diagnoses. Mol Psychiatry 23:1981–1989.
4. Stoddard J, Tseng W-L, Kim P, Chen G, Yi J, Donahue L, *et al.* (2017): Association of irritability and anxiety with the neural mechanisms of implicit face emotion processing in youths with psychopathology. JAMA Psychiatry 74:95.
5. Tseng W-L, Deveney CM, Stoddard J, Kircanski K, Frackman AE, Yi JY, *et al.* (2019): Brain mechanisms of attention orienting following frustration: Associations with irritability and age in youths. Am J Psychiatry 176:67–76.
6. Angold A, Costello EJ, Erkanli A (1999): Comorbidity. J Child Psychol Psychiatry 40:57–87.
7. Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TEJ, Glasser MF, *et al.* (2015): A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat Neurosci 18:1565–1567.
8. Wang H-T, Smallwood J, Mourao-Miranda J, Xia CH, Satterthwaite TD, Bassett DS, Bzdok D (2018): Finding the needle in high-dimensional haystack: A tutorial on canonical correlation analysis. ArXiv181202598 Cs Stat. Available at: http://arxiv.org/abs/1812.02598. Accessed June 23, 2020.
9. Xia CH, Ma Z, Ciric R, Gu S, Betzel RF, Kaczkurkin AN, *et al.* (2018): Linked dimensions of psychopathology and connectivity in functional brain networks. Nat Commun 9:3003.
10. Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, *et al.* (2017): Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med 23:28–38.
11. Dinga R, Schmaal L, Penninx BWJH, van Tol MJ, Veltman DJ, van Velzen L, *et al.* (2019): Evaluating the evidence for biotypes of depression: Methodological replication and extension of. Neuroimage Clin 22:101796.
12. Alexander LM, Escalera J, Ai L, Andreotti C, Febre K, Mangone A, *et al.* (2017): An open resource for transdiagnostic research in pediatric mental health and learning disorders. Sci Data 4:170181.
13. Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, *et al.* (2016): Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat Neurosci 19:1523–1536.
14. Winkler AM, Renaud O, Smith SM, Nichols TE (2020): Permutation inference for canonical correlation analysis. Neuromage 220:117065.
15. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, *et al.* (1997): Schedule for Affective Disorders and Schizophrenia for School-age Children—Present and Lifetime version (K-SADS-PL): Initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 36:980–988.
16. Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, Kaufman J, Neer SM (1997): The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. J Am Acad Child Adolesc Psychiatry 36:545–553.
17. Stringaris A, Goodman R, Ferdinando S, Razdan V, Muhrer E, Leibenluft E, Brotman MA (2012): The Affective Reactivity Index: A concise irritability scale for clinical and research settings: The Affective Reactivity Index. J Child Psychol Psychiatry 53:1109–1117.
18. Conners CK, Pitkanen J, Rzepa SR (2011): Conners 3rd Edition (Conners 3; Conners 2008). In: Kreutzer JS, DeLuca J, Caplan B, editors. Encyclopedia of Clinical Neuropsychology. New York, NY: Springer New York, 675–678.
19. Achenbach TM, Rescorla LA (2001): Manual for the ASEBA School-Age Forms & Profiles: An Integrated System of Multi-Informant Assessment. Burlington: University of Vermont Research Center for Children, Youth and Families.
20. Esteban O, Birman D, Schaer M, Koyejo OO, Poldrack RA, Gorgolewski KJ (2017): MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. PLoS One 12:e0184661.
21. 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.
22. Ciric R, Wolf DH, Power JD, Roalf DR, Baum GL, Ruparel K, *et al.* (2017): Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. NeuroImage 154:174–187.
23. Parkes L, Fulcher B, Yücel M, Fornito A (2018): An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. Neuroimage 171:415–436.
24. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, *et al.* (2018): Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb Cortex 28:3095–3114.
25. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, *et al.* (2002): Whole brain segmentation. Neuron 33:341–355.
26. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE (2014): Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 84:320–341.
27. Stewart D, Love W (1968): A general canonical correlation index. Psychol Bull 70:160–163.
28. Alnæs D, Kaufmann T, Marquand AF, Smith SM, Westlye LT (2019): Patterns of socio-cognitive stratification and perinatal risk in the child brain. Neuroscience 117:12419–12427.
29. Craske MG, Treanor M, Conway CC, Zbozinek T, Vervliet B (2014): Maximizing exposure therapy: An inhibitory learning approach. Behav Res Ther 58:10–23.
30. Hanlon CA, Dowdle LT, Henderson JS (2018): Modulating neural circuits with transcranial magnetic stimulation: Implications for addiction treatment development. Pharmacol Rev 70:661–683.
31. White LK, Sequeira S, Britton JC, Brotman MA, Gold AL, Berman E, *et al.* (2017): Complementary features of attention bias modification therapy and cognitive-behavioral therapy in pediatric anxiety disorders. Am J Psychiatry 174:775–784.
32. Pituch KA, Stevens JP (2016): Applied Multivariate Statistics for the Social Sciences, 6th ed. New York, NY: Routledge.
33. Tabachnick BG, Fidell LS (2001): Using Multivariate Statistics. Needham Heights, MA: Allyn and Bacon.

Shared and Anxiety-Specific Pediatric Psychopathology Dimensions Manifest Distributed Neural Correlates

Supplemental Information

Supplementary Methods

Detailed description of the methods

Participants

The present study comprised two independent samples. One sample was recruited at the National Institutes of Mental Health Intramural Research Program (NIMH-IRP) with the goal of investigating brain mechanisms mediating anxiety, irritability and disruptive behavior in youth. This sample was selected as the discovery sample because data was collected in-house on two identical scanners with the same imaging parameters. The other sample was recruited by the Child Mind Institute (CMI) as part of research assembling a publicly available, transdiagnostic sample of treatment-seeking youth (1) and served as the replication dataset. Imaging data of the replication dataset was collected in three different model scanners of varied field strengths.

Both samples comprised healthy volunteers (HV) and youth diagnosed with an anxiety disorder (ANX), disruptive mood dysregulation disorder (DMDD), or attention-deficit/hyperactivity disorder (ADHD). In both samples, diagnoses were established by licensed clinicians using the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS) (2). Exclusion criteria for both samples comprised neurological disorders, autism spectrum disorders, psychosis, bipolar disorder, substance use, MRI contraindications, and full-scale IQ < 70. Further, participants in each sample were required to have completed relevant questionnaire data. In the discovery sample, participants over age 18 and parents of minor participants gave written informed consent after receiving a complete description of the study; minors gave written assent. Procedures were approved by the Institutional Review Board of the NIMH.

Samples were similar in terms of sex ratios, proportions of anxiety disorders, oppositional defiant disorder, medication-free-to-medication-use ratios, as well as levels of parent-reported symptoms of irritability, disruptive behavior, and anxiety. However, the discovery sample was older, had a higher IQ, lower proportions of ADHD cases, higher proportions of diagnosis-free and DMDD cases, and lower levels of self-reported irritability and anxiety (Table S1**)**.

Table S1. Comparison of the Sample Characteristics

	Discovery sample	Replication sample	Statistic	p-value
Sex, male : female	1 : 1.1	1.2 : 1	$Chi^2_{(1)} = 3.4$	.07
Age in years, mean (SD, range)	13.8 (2.58, 9-19)	12.1 (2.83, 8-18)	$t_{(507)} = 6.72$	< .001
Intelligence score, mean (SD)	113.4 (11.16)	100.8 (15.49)	$t_{(474.2)} = 10.55$	< .001
Psychotropic medication, n : y	2.8 : 1	3.1 : 1	$Chi^2_{(1)} = 0.3$	.62
Diagnoses				
ADHD, %	27	53	$Chi^2_{(1)} = 30.8$	< .001
Anxiety disorder, %	44	39	$Chi^2_{(1)} = 1.2$	.26
DMDD, %	23	3	$Chi^2_{(1)} = 54.9$	< .001
HV, %	32	23	$Chi^2_{(1)} = 5.2$	.02
ODD, %	15	11	$Chi^2_{(1)} = 1.8$	.18
Questionnaires				
ARI-P, mean (SD)	2.8 (3.4)	3.1 (3.3)	U = 27291.0	.10
ARI-S, mean (SD)	2.6 (2.9)	3.3 (3.3)	U = 24910.5	< .01
SCARED-P, mean (SD)	15.4 (13.8)	15.2 (12.6)	U = 28788.0	.51
SCARED-S, mean (SD)	18.6 (15.0)	24.1 (16.4)	U = 23730.0	< .001
ADHD*, mean (SD)	.06 (4.93)	-.01 (3.47)	U = 29658.0	.91

Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; **ARI_P**, Affective Reactivity Index rated by the parent; **ARI_S**, Affective Reactivity Index rated by the child; **DMDD,** Disruptive Mood Dysregulation Disorder; **HV,** healthy volunteers; **ODD,** Oppositional Defiant Disorder; **SCARED-P,** Screen for Child Anxiety Related Disorders rated by the parent; **SCARED-C,** Screen for Child Anxiety Related Disorders rated by the child

* In the discovery sample, ADHD symptoms were rated by the parents using the Connors, whereas the same items were rated by the parents with the Child Behavior Checklist in the replication sample. Values presented here are normalized.

Further, we observed no differences between subsets of subjects from the discovery sample with data acquired on the two different 3T scanners (Table S2). In the replication dataset, sequence was confounded with the scan site. Thus, we compared the sample characteristics across the three sites. At the Staten Island, compared to the other two sites, we found that a significantly lower proportion of individuals met criteria for ADHD and a significantly higher proportion of individuals were classified as healthy volunteers. Further, youth-reported levels of irritability were lower at the CBIC site than the other two sites. The three sites were comparable with regard to all other sociodemographic and clinical variables (Table S3).

Table S2. Characteristics of the discovery sample separately for the two different scanners

	Scanner 1 (N=143)	Scanner 2 (N=40)	Statistic	p-value
Sex, male : female	1 : 1.1	1 : 1.2	$Chi^2_{(1)} = 0.08$	.78
Age in years, mean (SD, range)	13.9 (2.60)	13.6 (2.51)	$t_{(181)} = 0.61$	.54
Intelligence score, mean (SD)	113.5 (11.33)	113.0 (10.66)	$t_{(180)} = 0.29$	.77
Psychotropic medication, n : y	2.8 : 1	3.0 : 1	$Chi^2_{(1)} = 0.04$	.84
Diagnoses				
ADHD, %	31	25	$Chi^2_{(1)} = 0.62$	.43
Anxiety disorder, %	44	43	$Chi^2_{(1)} = 0.03$	.86
DMDD, %	21	25	$Chi^2_{(1)} = 0.30$	.59
HV, %	35	25	$Chi^2_{(1)} = 1.23$	.27
ODD, %	15	13	$Chi^2_{(1)} = 0.21$	.65
Questionnaires				
ARI-P, mean (SD)	2.9 (3.38)	2.7 (3.37)	U = 2757.0	.72
ARI-S, mean (SD)	2.5 (2.85)	2.7 (3.21)	U = 2853.0	.98
SCARED-P, mean (SD)	14.7 (13.87)	17.8 (13.41)	U = 2419.5	.14
SCARED-S, mean (SD)	18.4 (14.42)	19.3 (17.12)	U = 2798.5	.84
Conners-P, mean (SD)	4.7 (4.78)	5.0 (5.48)	U = 2775.0	.77

Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; **ARI_P**, Affective Reactivity Index rated by the parent; **ARI_S**, Affective Reactivity Index rated by the child; **Conners,** Conners III ADHD Rating Scales and specifically the sum score of the items that correspond to those in the ADHD subscale of the Child Behavior Checklist; **DMDD,** Disruptive Mood Dysregulation Disorder; **HV,** healthy volunteers; **ODD,** Oppositional Defiant Disorder; **SCARED-P,** Screen for Child Anxiety Related Disorders rated by the parent; **SCARED-C,** Screen for Child Anxiety Related Disorders rated by the child

Table S3. Characteristics of the replication sample separately for the three different sites/ scanners

	SI (N=112)	RUBIC (N=137)	CBIC (N=77)	Statistic	p-value
Sex, male : female	1.1 : 1	1.4 : 1	1.1 : 1	$\chi^2_{(2)} = 1.25$	.54
Age in years, mean (SD)	12.1 (2.89)	12.2 (2.72)	12.1 (2.83)	$F_{(2,323)} = 0.05$	.96
Intelligence score, mean (SD)	100.0 (14.91)	100.1 (15.62)	103.8 (15.99)	$F_{(2,323)} = 1.36$	.26
Psychotropic medication, y:n	1 : 4.6	1 : 3.0	1 : 2.1	$\chi^2_{(2)} = 5.35$	.07
Diagnoses					
ADHD, %	44	60	57	$\chi^2_{(2)} = 9.50$	.009
Anxiety disorder, %	35	37	46	$\chi^2_{(2)} = 2.94$	.23
DMDD, %*	5	1	1		.15
HV, %	35	17	17	$\chi^2_{(2)} = 13.4$	.001
ODD, %*	13	13	4		.08
Questionnaires					
ARI-P, mean (SD)	3.3 (3.40)	3.3 (3.36)	2.6 (2.99)	H = 1.77	.41
ARI-S, mean (SD)	3.6 (3.29)	3.7 (3.60)	2.3 (2.45)	H = 8.59	.01
SCARED-P, mean (SD)	15.8 (13.52)	16.3 (12.23)	12.5 (11.41)	H = 5.77	.06
SCARED-S, mean (SD)	25.6 (16.10)	23.9 (16.5)	22.1 (16.73)	H = 3.27	.20
CBCL ADHD, mean (SD)	4.96 (4.09)	5.95 (3.70)	5.42 (3.86)	H = 5.34	.07

Abbreviations: ADHD, Attention-Deficit/Hyperactivity Disorder; **ARI_P**, Affective Reactivity Index rated by the parent; **ARI_S**, Affective Reactivity Index rated by the child; **CBCL ADHD,** Attention-Deficit/ Hyperactivity subscale of the Child Behavior Checklist; **DMDD,** Disruptive Mood Dysregulation Disorder; **HV,** healthy volunteers; **ODD,** Oppositional Defiant Disorder; **SCARED-P,** Screen for Child Anxiety Related Disorders rated by the parent; **SCARED-C,** Screen for Child Anxiety Related Disorders rated by the child

* As cell sizes were <5, Fisher's Exact Test was calculated.

Clinical assessment

Anxiety symptoms were assessed with the Screen for Child Anxiety Related Disorders (SCARED) (3) completed by both parent and child. Its subscales of social anxiety, separation anxiety, panic, and school refusal were used as input for the CCA. Irritability symptoms were rated with the Affective Reactivity Index (ARI) by both parent and child (4). The first 6 items measuring intensity and frequency of angry mood and temper outbursts, which are rated on a scale from 0 to 2, were used as input for the CCA. In the discovery sample, disruptive behavior was quantified through seven items (failure to finish, difficulties concentrating, difficulties sitting still, impulsiveness, inattention, talking excessively and being loud) rated by the parent with the Conners (5); in the replication sample, these same seven items were rated on the Child Behavior Checklist (CBCL) (6). In both samples, all three dimensions were moderately correlated with one another (see Table S2).

Table S4. Associations among symptom dimensions. As correlation matrices are symmetric, we depict Spearman correlation coefficients (bold font) of the discovery dataset in the upper- right triangle and correlation coefficients of the replication sample in the lower-left triangle.

	ARI-P	ARI-Y	SCARED-P	SCARED-Y	ADHD
ARI-P		**.55**	**.32**	**.23**	**.50**
		5.1×10^{-16}	8.0×10^{-5}	.002	6.1×10^{-13}
ARI_Y	**.35**		**.38**	**.42**	**.28**
	4.5×10^{-11}		9.1×10^{-8}	2.6×10^{-9}	9.9×10^{-5}
SCARED-P	**.39**	**.18**		**.58**	**.23**
	4.2×10^{-13}	.001		1.5×10^{-17}	.002
SCARED-Y	**.10**	**.40**	**.26**		**.14**
	.074	4.2×10^{-14}	2.0×10^{-6}		.065
ADHD	**.45**	**.27**	**.26**	**.08**	
	3.0×10^{-17}	6.4×10^{-7}	2.0×10^{-6}	.177	

Abbreviations: ***ARI-P***, Affective Reactivity Index completed by the parent; ***ARI-Y***, Affective Reactivity Index completed by the child; ***SCARED-P***, Screen for Child Anxiety Related Disorders completed by the parent; ***SCARED-Y***, Screen for Child Anxiety Related Disorders completed by the child; ***ADHD***, 7 items assessing inattention and disruptive behavior taken from the Conners in the discovery dataset and from the Child Behavior Checklist in the replication dataset

Acquisition of imaging data

Imaging data of the discovery dataset were acquired on two identical 3.0 Tesla General Electric Signa scanners using a 32-channel head coil. Blood-oxygen-level-dependent (BOLD) changes during rest were measured for 10 minutes with a multi-echo planar imaging sequence (TR = 2000 msec, TE1/2/3 = 14.8 / 28.4 / 42.0msec, flip angle 77°, FoV 240x240x260, matrix size 64×64, 34 axial interleaved slices, slice thickness 3.8mm, bandwidth = 7812.5 Hz/Pixel). During this scan, participants were instructed to keep their eyes open and focus on white fixation cross in the middle of a black screen. In addition, a T1-weighted, magnetization-prepared, rapid-acquisition gradient echo (MPRAGE) sequence was acquired (TE = min full; TI = 425; FOV = 25.6; freq×phase = 256 × 256; flip angle = 7°; 1mm^3 voxels).

Imaging data of the replication dataset were acquired at three different sites (Staten Island – 1.5T Siemens Avanto: n=112, Rutgers University Brain Imaging Center – Siemens 3T Tim Trio: n=136, CitiGroup Cornell Brain Imaging Center – Siemens 3T Prisma: n=77). Scanning parameters have been described in detail elsewhere (1). Across sites, two different multiband echo planar imaging sequences were used to collect 10 minutes of resting-state data (Rutgers & CitiGroup: TR/ TE = 800/ 30 msec, flip angle 31°, 60 axial slices, resolution 2.4x2.4x2.4mm, multiband factor 6, sometimes administered as 1 and sometimes as 2 runs, 750 acquired volumes; Staten Island: TR/ TE = 1450/ 40 msec, flip angle 55°, 54 axial slices, resolution 2.5mm^3, multiband factor 3, 420 acquired volumes). Participants viewed a fixation cross on the center of a computer screen and were instructed to open or close their eyes at various points

throughout the scan. Further, two different MPRAGE sequences were acquired (Rutgers & CitiGroup: TR/ TE/ TI = 2500/ 3.15/ 1060 msec; 224 slices; resolution 0.8 mm^3, flip angle = 8°; Staten Island: TR/ TE/ TI = 2730/ 1.64/ 1000 msec; 176 slices; resolution 1 mm^3, flip angle = 7°).

Quality assessment of imaging data

Quality of the imaging data was assessed using MRIQC version 0.14.2 (7). Based on the indices provided by MRIQC (T1-weighted image: coefficient of joint variation, contrast to noise ratio, signal to noise ratios; fMRI: AFNI's outlier ratio, AFNI's quality index, DVARS, framewise displacement, signal to noise ratio, and temporal signal to noise ratio) and visual inspection, 50 participants from the discovery sample and 219 participants from the replication sample were excluded, mostly due to the presence of motion-related artefacts. Comparisons of the quality indices of the final discovery (N=182) and replication sample (N=326) showed large differences between the two samples (Figure S1). The quality of the functional data obtained at the two scanners in the discovery sample were comparable (see Table S5).

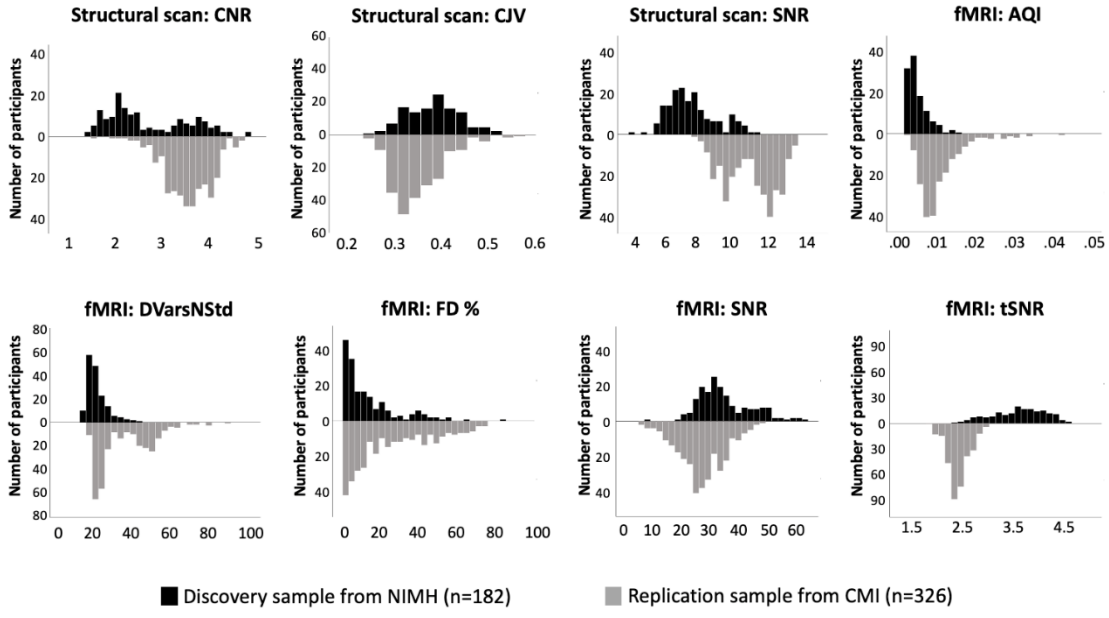


Figure S1. A selection of indices provided by MRIQC for the discovery (black) and replication samples (gray). Samples differ significantly in all indices (all p < .001).

Abbreviations: ***AQI***, AFNI's quality index (lower values are better); ***CJV***, coefficient of joint variation (lower values are better); ***CNR***, contrast to noise ratio (higher values are better); ***DVarsNStd***, per-image standard deviation of the temporal derivative of the data (lower values are better); ***FD%***, percent of timepoints above the framewise displacement threshold of 0.2 mm; ***SNR***, signal to noise ratio (higher values are better); ***tSNR***, temporal signal to noise ratio (higher values are better).

Table S5. rsfMRI quality indices separately for the two different scanners used for the discovery sample

	Scanner 1 (N=143)	Scanner 2 (N=40)	Statistic	p-value
aor	.006 (.006)	.008 (.007)	$t_{(181)} = -1.69$	.094
aqi	.005 (.003)	.006 (.003)	$t_{(181)} = -1.60$	.112
dvars_nstd	19.5 (5.62)	17.9 (4.86)	$t_{(181)} = 1.62$	.106
FD, mean	.14 (.108)	.17 (.120)	$t_{(181)} = -1.65$	.101
FD, percent	13.2 (15.91)	16.8 (14.57)	$t_{(181)} = -1.28$	.202
tsnr	62.78	60.24	$t_{(181)} = 1.04$	.301

Abbreviations: aor, AFNI's outlier ratio; **aqi**, AFNI's quality index; **dvars_nstd**, the per-image standard deviation of the temporal derivative of the data. A measure of how much the intensity of a brain image changes in comparison to the previous timepoint (as opposed to the global signal); **FD,** framewise displacement; **tsnr,** temporal signal-to-noise ratio

Image processing

Based on MRIQC (v0.14.2), four initial volumes from the functional data from the discovery sample and five initial volumes from each run of the replication sample were discarded. Image processing used the FMRIPREP (v1.3.2), (8) an automated processing pipeline based on Nipype (9). First a reference image and mask were created based on the initial functional volumes. Next, functional data were slice time corrected using 3dTshift from AFNI, motion corrected using MCFLIRT from FSL (10), and co-registered to the structural volume using boundary-based registration with 9 degrees of freedom (11); these spatial transformations were concatenated and applied all in a single interpolation step. Finally, ICA-based Automatic Removal of Motion Artifacts (AROMA) was used to non-aggressively denoise the time courses (14). We refrained from motion scrubbing as it has been repeatedly shown that AROMA performs adequately in reducing motion-related artefacts. (15,16) In addition, the use of AROMA allows us to keep a constant number of timepoints across participants and does not compromise the temporal dynamics, which facilitate approaches to data analysis. For the multi-echo data from the discovery dataset the Tedana T2* workflow (17) was used to create an optimally weighted combination of the three echo times.

Structural scans were subjected to FreeSurfer processing, as called from FMRIPREP. The T1-weighted images were initially corrected for magnetic field inhomogeneities and skull-stripped (18). A mass of segmented white matter voxels was produced for each hemisphere, and a mesh of triangular faces was tightly built around this mass, using two triangles per exposed voxel face. The mesh was smoothed taking into account the local intensity in the original images, (19) at a subvoxel resolution. Topological defects were corrected (20,21) ensuring that the surfaces for each hemisphere had the same topological properties of a sphere. A second iteration of smoothing was applied, resulting in a fiducial representation of the interface

between gray and white matter (the white surface). The external cortical surface (the pial surface) was produced by nudging outwards the white surface towards a point where the tissue contrast was maximal, maintaining constraints on smoothness and on the possibility of self-intersection. (22) Once the cortical models were complete, surfaces were homomorphically transformed into a sphere, registered and resampled to a common space ("fsaverage") based on individual cortical folding patterns so as to match cortical geometry across subjects. (23) Functional, processed resting stated time series were then resampled from voxel space to surface space.

Network construction

Because the shortest time series contained 300 time points, a published 200-region parcellation scheme (24) was used to define the network nodes. This generated for data analysis more timepoints than parcels. The parcellation scheme used was originally derived through the combination of two algorithms (boundary mapping and clustering), such that parcels are expected to be functionally and connectionally more homogeneous than, for example, random parcels. These parcels have previously been assigned to known intrinsic functional networks. (24,25) The 200 cortical areas were augmented by the inclusion of 8 subcortical regions per hemisphere obtained using FreeSurfer segmentation; these were nucleus accumbens, nucleus caudatus, pallidum, putamen, amygdala, hippocampus, thalamus and ventral diencephalon (a miscellaneous group of structures that include hypothalamus, mammillary bodies, subthalamic nuclei, substantia nigra, red nucleus, and medial and lateral geniculate nuclei; white matter areas such as the zona incerta, crus cerebri, lenticular fasciculus, and medial lemniscus are also included in this area). Thus, the functional connectivity network comprised 216 nodes. Framewise displacement and spatial standard deviation of the temporal difference data (dvars) were regressed out from the time series; global signal was not regressed out or used as a nuisance variable. Functional connectivity was quantified using partial correlations, which offer an estimate of direct (as opposed to indirect or shared) connectivity between each pair of nodes; these correlations were then transformed to z-scores using Fisher's transformation, i.e., z = arctanh(r). As the resulting network matrices are symmetric, we only analyzed the 23220 edges on one side of the diagonal of the connectivity matrix (i.e., 216×215/2 connections). For the participants of the discovery sample that had two runs, z-scores for each edge were averaged and the variance was re-scaled by a factor of $\sqrt{2}$.

CCA, ICA, and permutation testing

Covariates were regressed out from the imaging and clinical variables before dimensionality reduction. These covariates comprised age, sex, race, IQ, psychotropic medication, and scanner for the discovery sample; plus site and sequence for the replication sample. An intercept was also included as a nuisance variable, thus obviating the need for further mean-centering. Next, while we included all 29 symptom ratings, dimensionality of rsfMRI was reduced using principal component analysis (PCA) prior to CCA.(26) Residuals were then projected, after dimensionality reduction, to a lower dimensional space where data are exchangeable, thus mitigating spurious dependencies among observations introduced by residualization (27).

CCA is a multivariate method that aims to reduce the correlation structure between two sets of variables (here rsfMRI and symptoms). In this approach, the simplest possible form is sought through linear transformations of the variables within each set. That is, given two sets of variables (here, symptom ratings, X, and imaging data, Y), the method seeks linear mixtures (i.e., canonical variables, CVs) within each set (i.e., $U = A \times Y$ and $V = B \times X$), such that each resulting mixture (U) from one set is maximally correlated with a corresponding mixture (V) from the other set, but uncorrelated with all other mixtures in either set. As the number of canonical variables is determined by the smallest input dataset, we obtained 29 CVs.

Given the CCA orthogonality constraint, it is possible, for canonical correlations of similar magnitudes, for small perturbations within a data set to generate solutions that are arbitrarily rotated. To mitigate this problem, aid with interpretation, and use principles adopted in previous studies (28–30) we subjected the results of CCA to Independent Component Analysis (ICA) (31,32). Unlike in previous studies, however, we did not use the canonical loadings to produce a single set of independent components for both clinical and imaging data, to which canonical variables would be correlated. Rather, we subjected the canonical variables from both sides to a joint ICA after concatenation of U and V, nesting the whole procedure into the iterative algorithm used for permutation inference for CCA, described below. Thus, rather than simply using ICA to characterize CCA results, we tested the significance of the joint CCA+ICA correlations between underlying latent factors that associate imaging and clinical data. The permutation test involved, for each permutation, re-estimating the post-ICA canonical correlations, while removing the variance already explained by all previous canonical variables, i.e., the ones with stronger canonical correlations than the current one, in an iterative procedure (27); in other words, for each permutation, an iterative re-estimation of the canonical correlations is performed, in which only variance not already explained by previous canonical

variables is used. Canonical correlations were considered significant at alpha = 0.05 after 10000 permutations and when using family wise error rate (FWER) correction, with a closed testing procedure. We also report the non-symmetric redundancy index (33), which gives the mean variance of the clinical data explained by imaging data, and vice versa.

Code related to this analysis is available at https://github.com/JuliaLinke/Linke_jointCCAICA.

Similarities across scanners

In our analysis, we adjusted for site and sequence using a linear model. The model includes one regressor per site (scanner) such that non-linear effects across sites are naturally accommodated. Nonetheless, differences in the distributions of the data across scanners/ sites may exist. Thus, we compared the distributions of the residuals stratified by site, using a two-sample Kolmogorov-Smirnov (KS) test, comparing two sites at a time. We note, however, that CCA is invariant to linear transformations of the input data, which in turn allows infinite possibilities of inputs for this KS test. Since exchangeability is asserted at the level of the canonical variables, we subject these to the KS-test. That is, we break the canonical variables per scanner in the discovery sample, and per site/scanner in the replication sample, and investigate distributional differences. Because this analysis uses the same data used in the CCA+ICA inference, it is already free of all nuisance variables (including age, sex, comorbidities, etc.), such that eventual residual differences would be the ones impacting CCA+ICA results. The results of this analysis are shown in Supplementary Table S6. The significant canonical variates were not affected by differences in the residuals of the distributions of the input data.

Table S6. Comparison of the distributions of the residuals stratified by scanner and site. Gray background indicates the significant canonical variates.

	Discovery sample						Replication sample					
CV	**Scanner 1 vs. Scanner2**			**Site 1 vs. Site 2**			**Site 1 vs. Site 3**			**Site 2 vs. Site 3**		
	KS	**p_{uncorr}**	**p_{FWER}**	**KS**	**p_{uncorr}**	**p_{FWER}**	**KS**	**p_{uncorr}**	**p_{FWER}**	**KS**	**p_{uncorr}**	**p_{FWER}**
1	0.18	0.18	0.18	0.19	0.19	0.19	0.22	0.22	0.22	0.1	0.1	0.1
2	0.21	0.26	1	0.2	0.06	1	0.13	0.01	0.38	0.12	0.65	1
3	0.32	0.13	0.98	0.22	0.05	0.99	0.21	0.26	1	0.17	0.49	1
4	0.25	0	0.08	0.22	0.02	0.76	0.17	0.01	0.45	0.11	0.11	1
5	0.17	0.03	0.63	0.18	0.02	0.84	0.17	0.06	1	0.06	0.52	1
6	0.17	0.32	1	0.28	0.1	1	0.23	0.05	0.99	0.1	0.99	1
7	0.29	0.29	1	0.24	0	0.1	0.11	0	0.22	0.16	0.65	1
8	0.36	0.01	0.22	0.23	0.01	0.47	0.16	0.37	1	0.1	0.14	1
9	0.16	0	0.01	0.21	0.01	0.61	0.16	0.08	1	0.11	0.72	1
10	0.23	0.41	1	0.17	0.04	0.96	0.22	0.07	1	0.2	0.53	1
11	0.18	0.07	0.89	0.23	0.14	1	0.15	0	0.33	0.13	0.03	0.95
12	0.24	0.26	1	0.19	0.02	0.75	0.14	0.1	1	0.08	0.34	1
13	0.14	0.06	0.81	0.22	0.07	1	0.18	0.15	1	0.12	0.9	1
14	0.33	0.55	1	0.19	0.02	0.88	0.13	0.03	0.93	0.1	0.47	1
15	0.2	0	0.04	0.2	0.07	1	0.16	0.24	1	0.08	0.69	1
16	0.12	0.14	0.99	0.18	0.05	0.98	0.17	0.09	1	0.11	0.87	1
17	0.13	0.73	1	0.14	0.09	1	0.18	0.05	0.99	0.17	0.58	1
18	0.36	0.62	1	0.16	0.29	1	0.12	0.04	0.96	0.12	0.1	1
19	0.24	0	0.01	0.18	0.16	1	0.14	0.35	1	0.06	0.49	1
20	0.11	0.05	0.79	0.17	0.09	1	0.22	0.19	1	0.13	0.99	1
21	0.19	0.79	1	0.19	0.14	1	0.12	0	0.34	0.1	0.33	1
22	0.18	0.19	1	0.15	0.07	1	0.16	0.29	1	0.09	0.69	1
23	0.15	0.26	1	0.18	0.22	1	0.11	0.09	1	0.12	0.81	1
24	0.14	0.47	1	0.19	0.1	1	0.17	0.46	1	0.09	0.45	1
25	0.15	0.58	1	0.17	0.08	1	0.12	0.05	0.98	0.1	0.79	1
26	0.33	0.45	1	0.28	0.14	1	0.22	0.28	1	0.12	0.65	1
27	0.12	0	0.05	0.18	0	0.08	0.17	0	0.25	0.11	0.4	1
28	0.11	0.73	1	0.23	0.08	1	0.12	0.05	0.98	0.16	0.54	1
29	0.09	0.85	1	0.19	0.01	0.69	0.15	0.33	1	0.17	0.17	1

Abbreviations: CV, canonical variate; **KS,** Kolmogorov-Smirnoff test statistic

Replicability

Robustness of the CVs was determined based on three criteria: (1) stability within the same dataset across varying input-to-participant ratios, (2) similarities of latent clinical patterns across the two independent samples, and (3) similarities of latent connectivity patterns across the two independent samples. A prior CCA study in youth reported replicability only for clinical but not rsfMRI patterns (34). Thus, we decided to evaluate the replicability of the clinical and the connectivity patterns as separate criteria.

Few studies exist to guide the decision of how many principal components should be used as input for a CCA analysis, given a specific sample size. Therefore, we performed three analyses that varied in the input-to-participant ratios within each sample. The primary analyses used an input-to-participant ratio of 1:2, which translated into 64 rsfMRI components explaining 75% of the between-subject variance in rsfMRI-connectivity in the discovery sample. In the replication cohort, 134 rsfMRI components were used; these explained only 57% of the variance, possibly due to more heterogeneous imaging parameters. This primary analysis was supplemented by two secondary analyses using input-to-participant ratios of 1:3 and 1:4 to reduce risks of overfitting, at the expense of explaining less variance. In the discovery sample, 32 and 17 components, which explained 60% and 47% of the between-subject connectivity variance, respectively, were used in these additional analyses. In the replication dataset, 79 and 53 components that explained 55% and 39% of the between-subject variance, respectively, entered these analyses. Results were compared across the three ratios by examining cross-correlations among CCA components (e.g., corr($v_{1|1:2}$, $Y_D \times a_{1|1:3}$) and corr($u_{1|1:2}$, $X_D \times b_{1|1:3}$)). Statistical significance was determined using 10000 permutations, with a threshold of $p_{FWER} < .05$ within each set of comparisons. As psychiatric symptoms might relate to components that explain relatively little variance in the imaging data, we will also discuss CVs that solely replicated at the 1:3 ratio but could be found in the replication cohort.

To test the second and third criteria, we used joint CCA+ICA in the replication dataset. Canonical weights from each dataset were applied to the input data from the other dataset; these products were then correlated with the CVs identified in that dataset (e.g., corr($v_{1|D}$, $Y_D \times a_{1|R}$) and corr($u_{1|D}$, $X_D \times b_{1|R}$)). Clinical and connectivity patterns were considered replicable when both the application of weights from the discovery to the replication dataset and the application of weights from the replication to discovery dataset yielded statistically significant associations. We used 10000 permutations to establish significance. However, thresholds differed for clinical and connectivity patterns. We used a stringent threshold of $p_{FWER} < .05$ to determine replicability of the imaging and clinical patterns; additionally, we also investigate a more lenient threshold of $p_{uncorr} < .05$ for replicability of the connectivity pattern. This decision was motivated by two factors. First, a prior CCA study finding replicable clinical patterns did not report replicable connectivity patterns across two subsets of a single sample (9). This raises questions as to whether any evidence of replicability can be detected with even liberal statistical thresholds. Second, in the current study, significant differences exist between cohorts in all metrics quantifying the quality of the imaging data (Figure S1); this contrasts with the broadly similar profiles for symptom ratings (Table S1 and Figure S1).

Interpretation of canonical variables

CVs might be interpreted by examining (a) the standardized coefficients of the CVs (canonical weights), (b) the correlations between the CVs and the input data (canonical loadings), and (c) the amount of variance explained (redundancy index). (33) Since input variables were moderately correlated, interpretations discuss canonical loadings, as multicollinearity might result in low canonical weights. Prior literature (26,30) highlights loadings with an absolute value larger than 0.20 in either positive or negative directions. While we present all loadings, discussions focus on loadings above the 0.20 threshold, despite the arbitrary nature of this threshold. In fact, we extend this approach by limiting our focus to replicating clinical loadings, i.e., loadings $|r| > 0.2$ that could be observed across samples. Similarly, we emphasize edge loadings that replicate across samples. However, given the differences in the quality of the imaging data across samples, we apply a more lenient threshold of $|r| > 0.15$ to the replication cohort.

Additional Results

Redundancy index

The redundancy index was calculated as in (33). In the discovery sample, CVs explained 14.2% of the variance of the symptom ratings from rsfMRI-connectivity, though only 2.2% of the between-subject variance in brain connectivity was explained by the clinical CVs. In the replication cohort, 3.4% of the variance of the symptom ratings was explained by the significant CVs from resting-state connectivity, whereas only 0.4% of the between-subject variance in brain connectivity was explained by the significant CVs from symptom ratings.

Canonical correlations in the discovery dataset

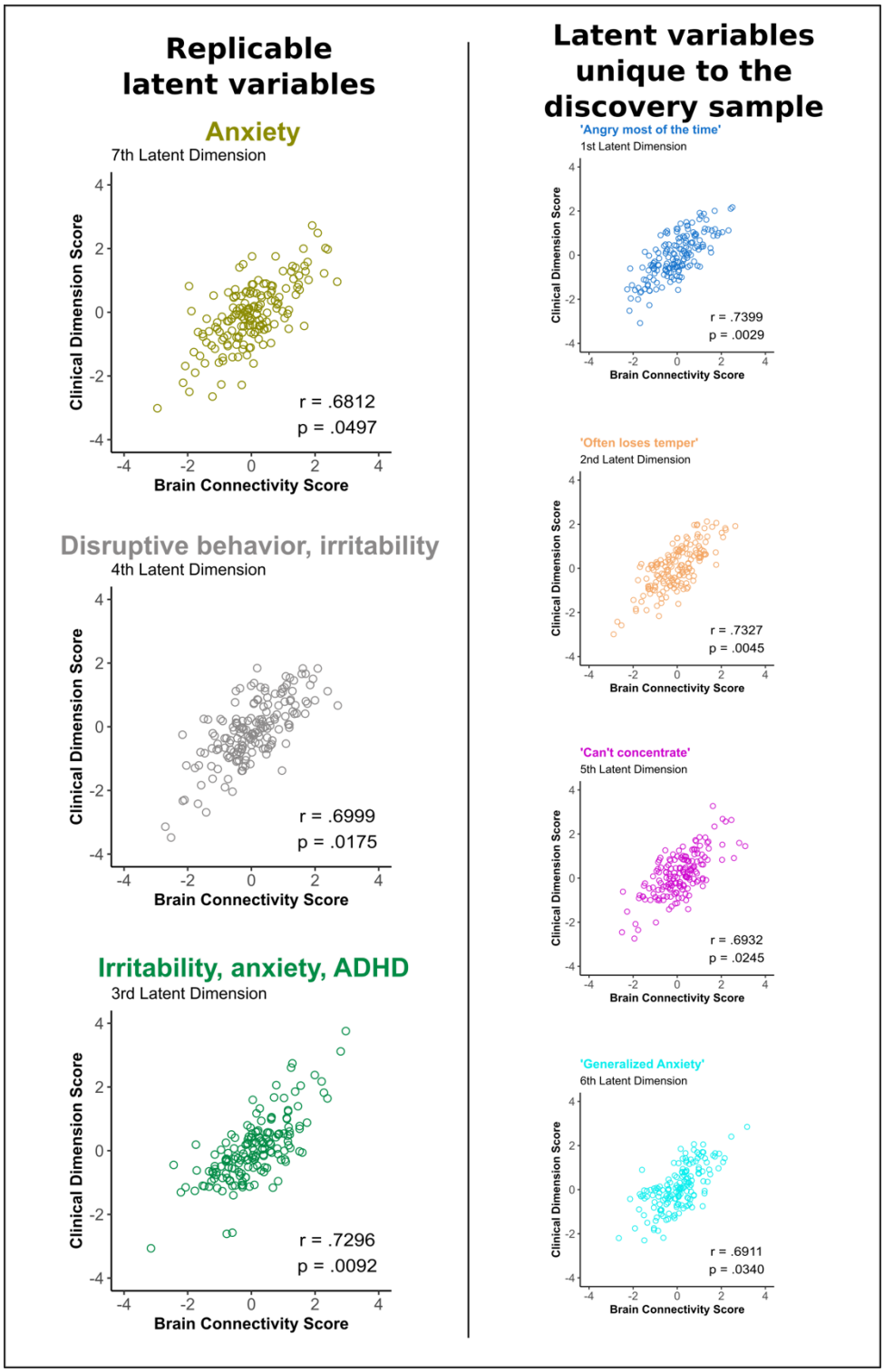
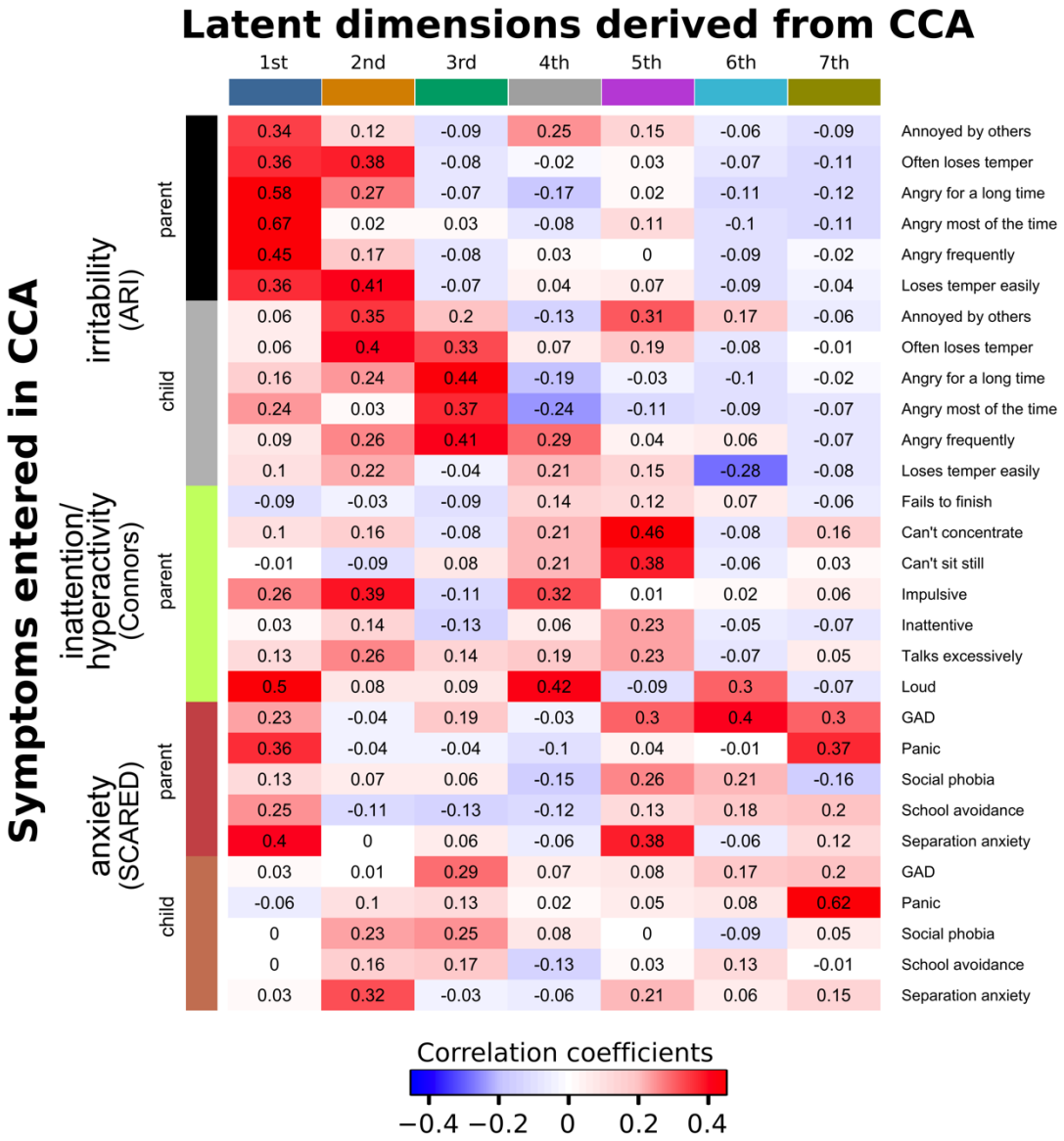


Figure S2. Illustration of the seven significant canonical variates in the discovery dataset. Scatter plots show the latent dimensions (linear combinations of brain connectivity scores obtained during rsfMRI in the horizontal axis, and linear combinations of clinical scores derived from symptom ratings in the vertical axis). For each latent dimension unique to the discovery dataset the highest loading clinical item is included. P-values are corrected for multiple testing using the family wise error rate.

Clinical loadings in the discovery dataset

Table S7. Clinical loadings in the discovery dataset



Edge loadings in the discovery data set

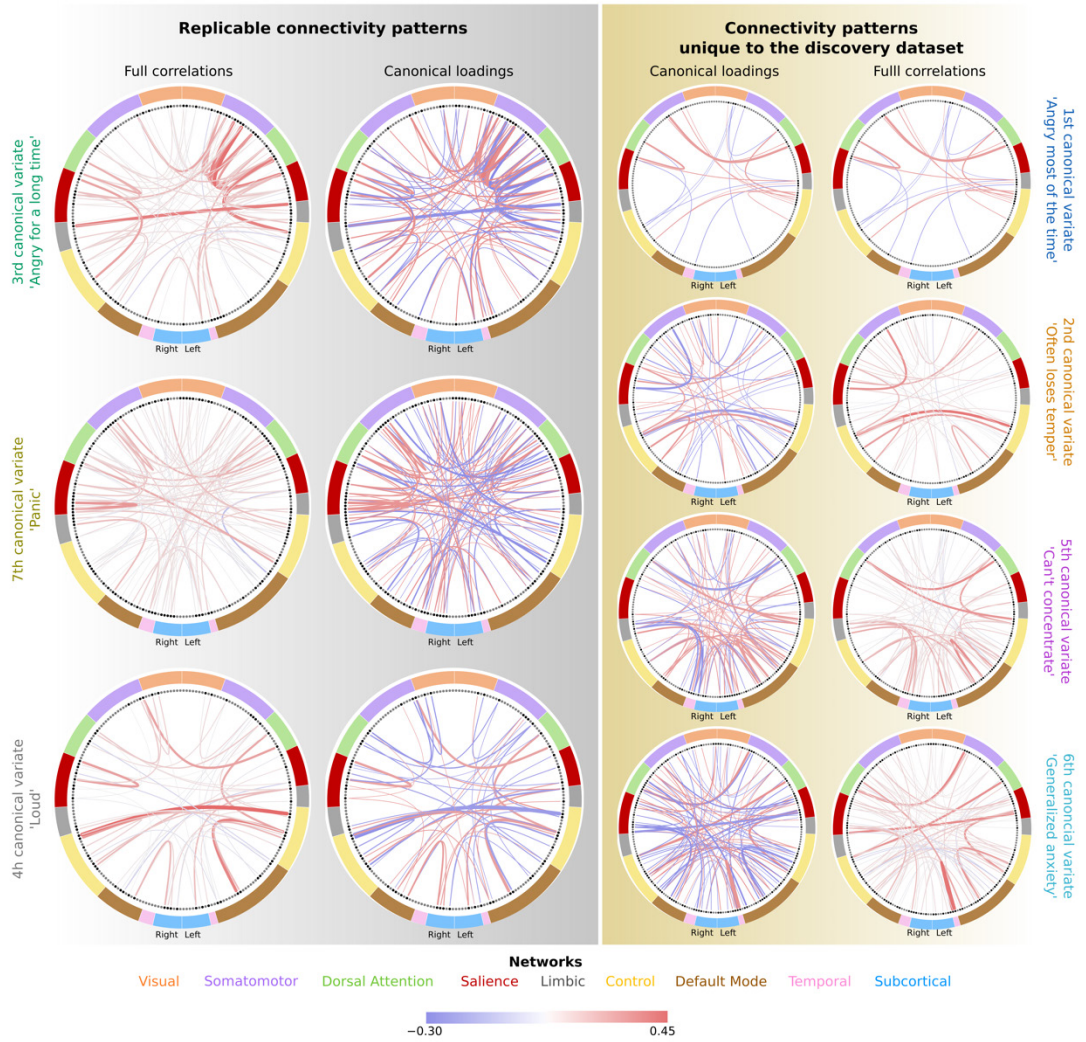


Figure S3. Overview of connectivity patterns associated with the three replicating latent variables. The two inner columns show the associations between the latent dimensions and resting-state connectivity patterns thresholded at |r| > 0.2. Red colored lines represent positive associations between the latent dimension and the edge, whereas blue lines indicate a negative relationship. The intensity of the lines shows the magnitude of the correlation on a scale from -1 to 1, and the thickness reflects the importance of the connection in general (i.e., full correlations). The two outer columns indicate the strength (thicker lines represent stronger correlations) and direction (red indicates positive and blue resembles negative connections) of the full correlations between the relevant nodes. For each hemisphere, the colors on the outer rings represent the nine large-scale networks our parcellation scheme maps onto: visual, somatomotor, dorsal attention, salience, limbic, control, default mode, temporal, and the eight subcortical structures taking from the FreeSurfer segmentation.

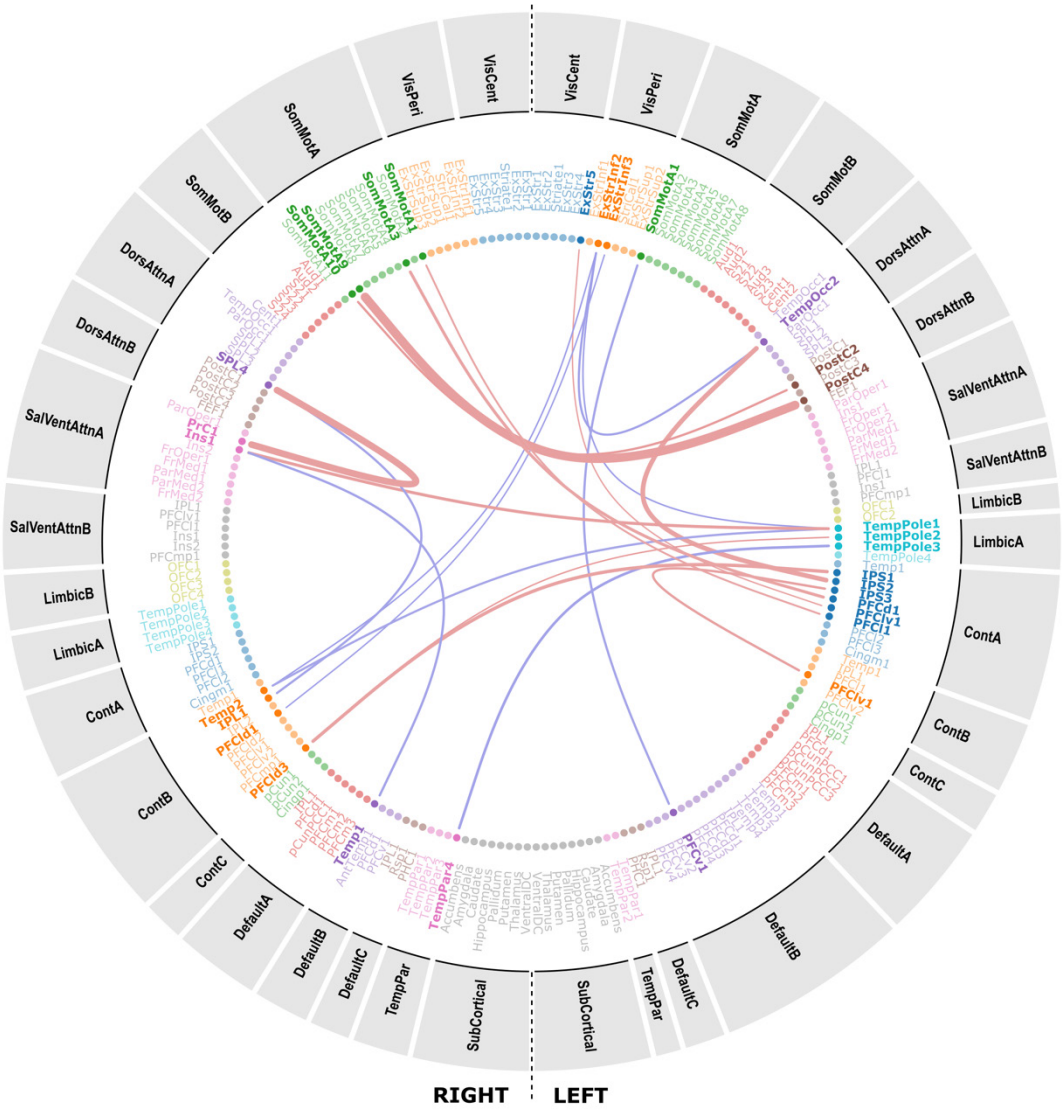


Figure S4. Enlarged depiction of the connectivity pattern associated with the 1st canonical variate in the discovery dataset. Edges are thresholded at |r| > 0.2.

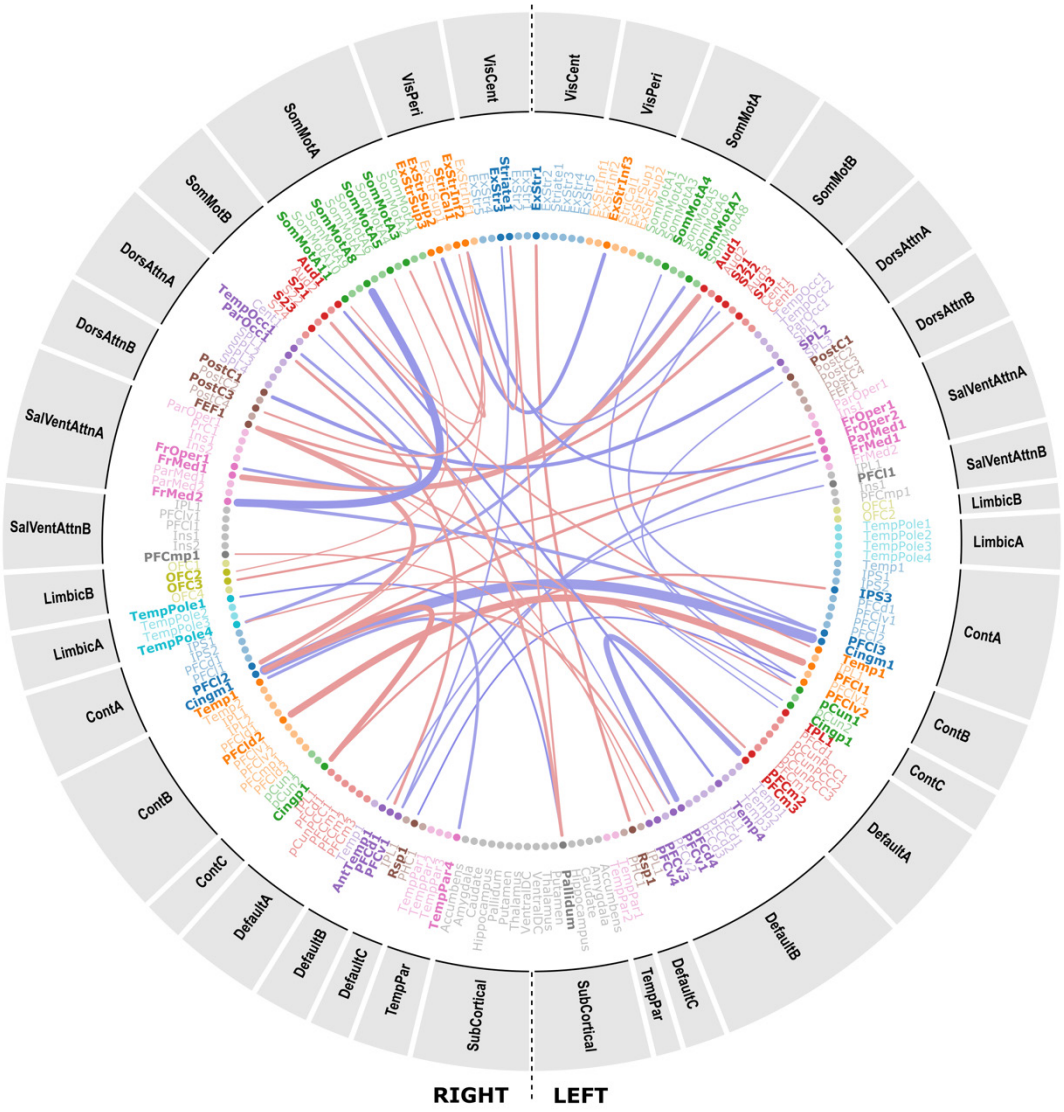


Figure S5. Enlarged depiction of the connectivity pattern associated with the 2nd canonical variate in the discovery dataset. Edges are thresholded at |r| > 0.2.

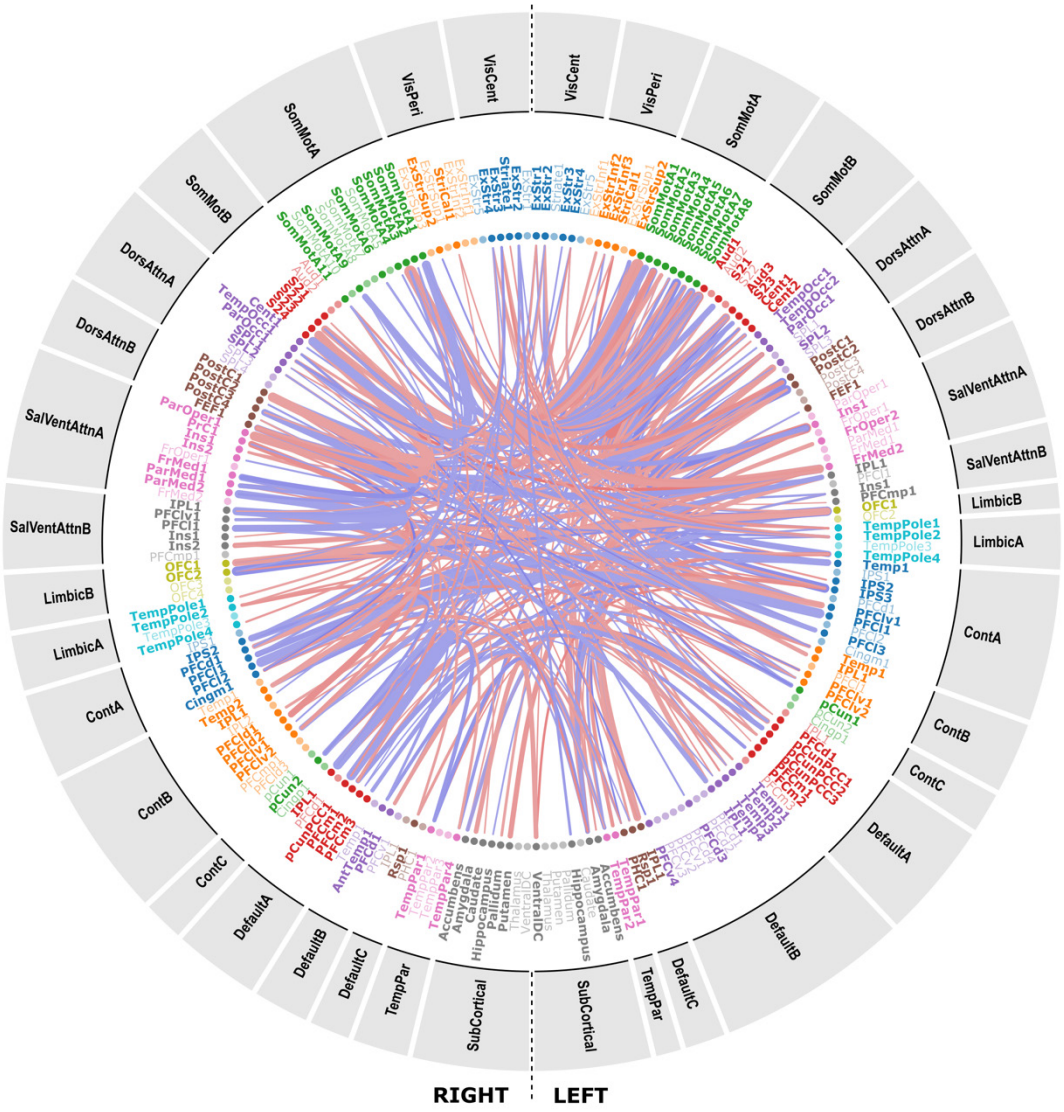


Figure S6. Enlarged depiction of the connectivity pattern associated with the 3rd canonical variate in the discovery dataset. Edges are thresholded at |r| > 0.2.

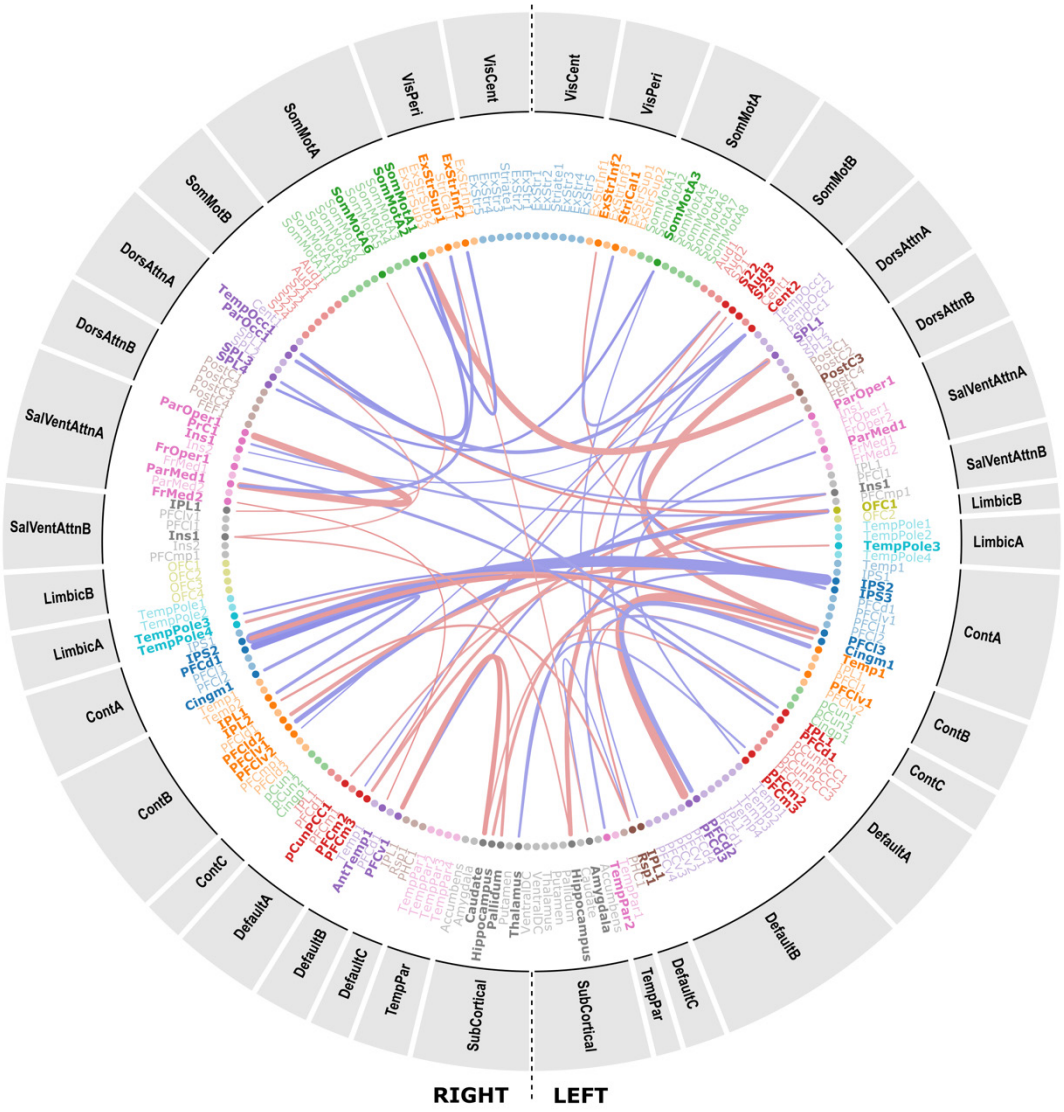


Figure S7. Enlarged depiction of the connectivity pattern associated with the 4th canonical variate in the discovery dataset. Edges are thresholded at |r| > 0.2.

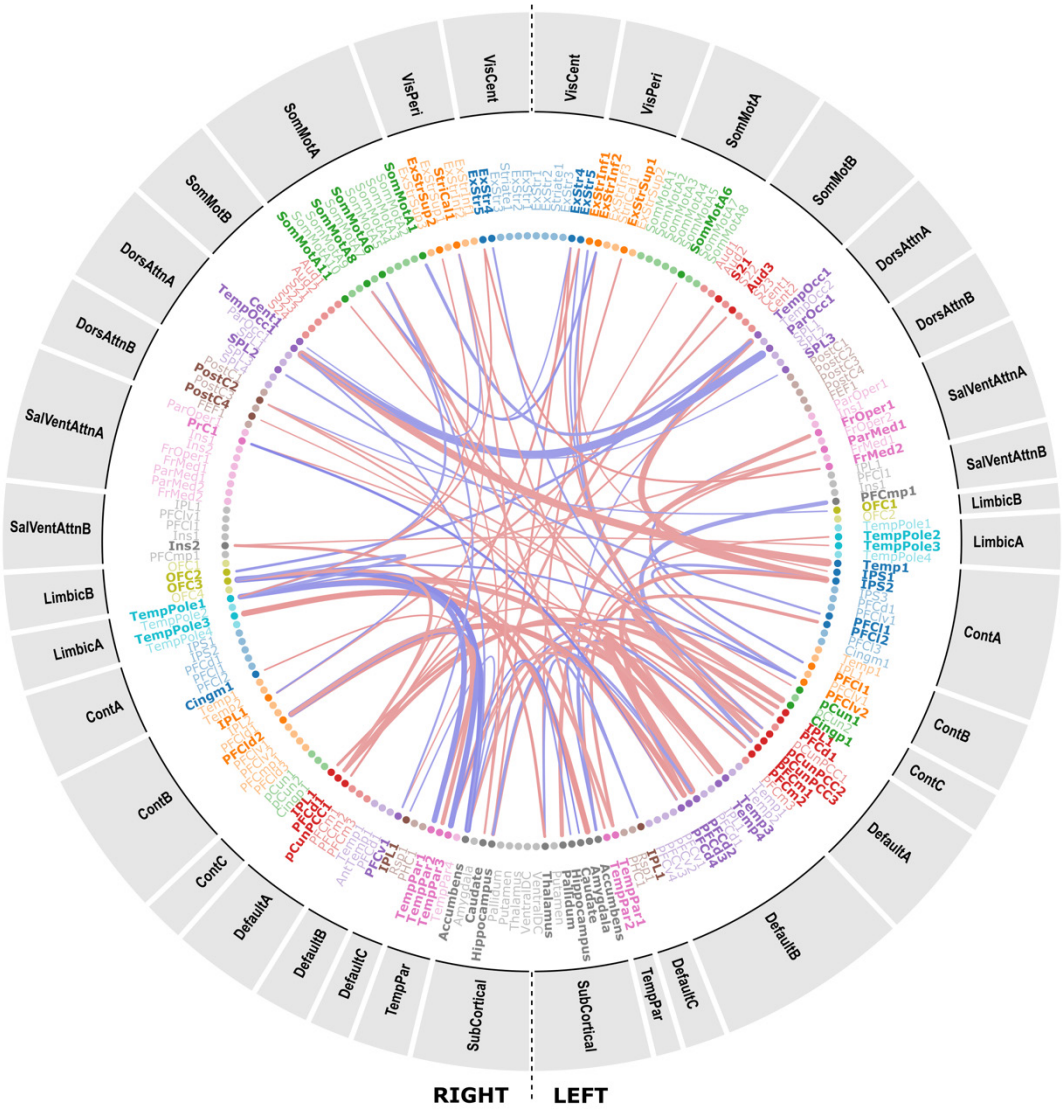


Figure S8. Enlarged depiction of the connectivity pattern associated with the 5th canonical variate in the discovery dataset. Edges are thresholded at |r| > 0.2.

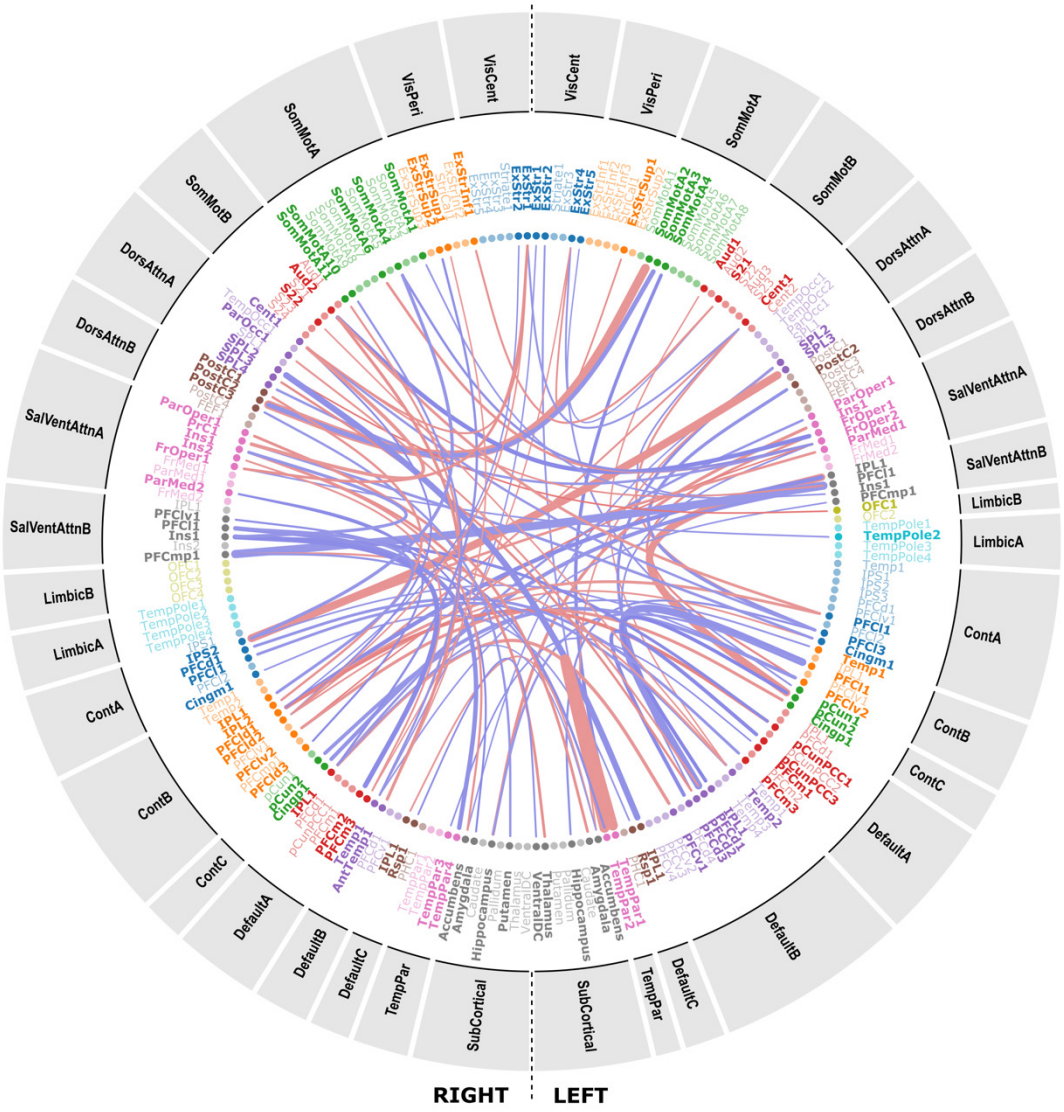


Figure S9. Enlarged depiction of the connectivity pattern associated with the 6th canonical variate in the discovery dataset. Edges are thresholded at |r| > 0.2.

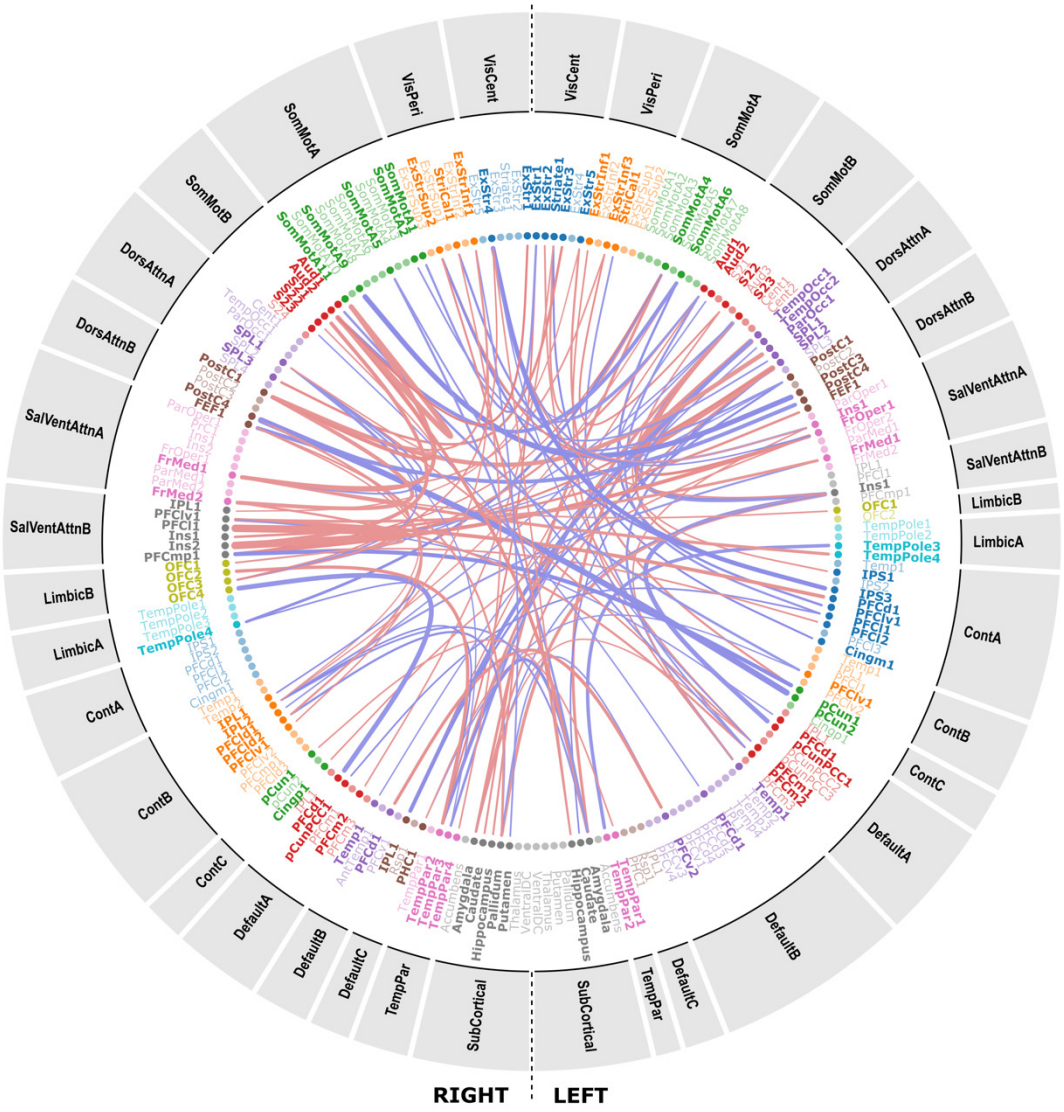


Figure S10. Enlarged depiction of the connectivity pattern associated with the 7th canonical variate in the discovery dataset. Edges are thresholded at |r| > 0.2.

Stability of results in the discovery dataset

When applying a input-to-participant ratio of 1:3 to the to select the PCA-components that entered the CCA, CCA yielded 17 significant CVs (CV_1: r = .59, p = 0.0002; CV_2: r = .53, p = 0.0002; CV_3: r =.52 , p = 0.0002 ; CV_4: r = .50, p = 0.0003; CV_5: r = .50, p = 0.0003; CV_6: r = .49, p = 0.0003; CV_7: r = .49, p = 0.0003; CV_8: r = .49, p = 0.0003; CV_9: r = .48, p = 0.0003; CV_{10}: r = .47, p = 0.0007; CV_{11}: r = .46, p = 0.0010; CV_{12}: r = .45, p = 0.0015; CV_{13}: r = .45, p = 0.0021; CV_{14}: r = .45, p = 0.0034; CV_{15}: r = .45, p = 0.0078; CV_{16}: r = .44, p = 0.0188; CV_{17}: r = .42, p = 0.0485). When applying an input-to-participant ratio of 1:4 to the to select the PCA-components that entered the CCA, CCA yielded 4 significant CVs (CV_1: r = .59, p = 0.0031; CV_2: r = .52, p = 0.0127; CV_3: r =.51, p = 0.0218; CV_4: r = .51, p = 0.0324).

Table S8. Cross-loadings of the clinical side for the 2:1 and 3:1 participant-to-input ratios in the discovery dataset. P-values were obtained using 10000 permutations corrected over 119 tests. Colored cells indicate stability on both the clinical and imaging side. Light green and yellow cells further indicate the CVs highlighted in the main text, which replicate across samples for both clinical and imaging loadings in green or only clinical loadings in purple.

	$V_{1\|D1:2}$	$V_{2\|D1:2}$	$V_{3\|D1:2}$	$V_{4\|D1:2}$	$V_{5\|D1:2}$	$V_{6\|D1:2}$	$V_{7\|D1:2}$
$V_{1\|D1:3}$	**-0.07** .3946/ 1	**-0.04** .6065/ 1	**0.97** .0001/ .0001	**0.02** .8042/ 1	**0.04** .6328/ 1	**0.01** .9461/ 1	**-0.01** .8763/ 1
$V_{2\|D1:3}$	**-0.13** .0983/ 1	**-0.91** .0001/ .0001	**-0.06** .4244/ 1	**0.06** .4448/ 1	**0.02** .8338/ 1	**0.14** .0688/ .9999	**-0.04** .5826/ 1
$V_{3\|D1:3}$	**-0.24** .0021/ .2248	**-0.05** .4987/ 1	**0.05** .4866/ 1	**0.08** .2816/ 1	**0.00** .9927/ 1	**-0.06** .4743/ 1	**-0.04** .5954/ 1
$V_{4\|D1:3}$	**0.01** .8666/ 1	**0.03** .7147/ 1	**0.04** .6477/ 1	**-0.16** .0368/ .9915	**0.02** .8453/ 1	**0.00** .9981/ 1	**-0.03** .6672/ 1
$V_{5\|D1:3}$	**-0.05** .4887/ 1	**-0.05** .5111/ 1	**-0.05** .5264/ 1	**-0.07** .3556/ 1	**0.05** .5558/ 1	**-0.09** .2776/ 1	**0.00** .9500/ 1
$V_{6\|D1:3}$	**-0.12** .1339/ 1	**0.06** .4584/ 1	**-0.04** .6109/ 1	**0.15** .0604/ .9995	**0.91** .0001/ .0001	**0.04** .6218/ 1	**0.00** .9682/ 1
$V_{7\|D1:3}$	**-0.64** .0001/ .0001	**0.12** .1228/ 1	**-0.09** .2315/ 1	**-0.37** .0001/ .0002	**-0.04** .5919/ 1	**0.01** .9327/ 1	**0.11** .1638/ 1
$V_{8\|D1:3}$	**-0.17** .0253/ .9548	**0.00** .9516/ 1	**0.03** .7016/ 1	**-0.21** .0087/ .6303	**-0.01** .9396/ 1	**0.09** .2467/ 1	**0.02** .7522/ 1
$V_{9\|D1:3}$	**0.14** .0814/ 1	**-0.02** .7576/ 1	**0.06** .4563/ 1	**-0.22** .0050/ .4684	**-0.01** .9054/ 1	**-0.09** .2590/ 1	**0.01** .8824/ 1
$V_{10\|D1:3}$	**0.01** .9370/ 1	**0.07** .3890/ 1	**-0.04** .6269/ 1	**0.14** .0704/ 1	**-0.07** .3969/ 1	**-0.05** .4894/ 1	**-0.86** .0001/ .0001
$V_{11\|D1:3}$	**-0.19** .0172/ .8602	**0.02** .7941/ 1	**-0.05** .4919/ 1	**0.43** .0001/ .0001	**-0.08** .2930/ 1	**-0.03** .7273/ 1	**0.06** .4311/ 1
$V_{12\|D1:3}$	**0.02** .7860/ 1	**0.04** .5929/ 1	**0.00** .9814/ 1	**0.16** .0354/ .9889	**-0.06** .4499/ 1	**0.01** .9170/ 1	**0.03** .7224/ 1
$V_{13\|D1:3}$	**0.00** .9791/ 1	**0.03** .6665/ 1	**0.04** .5820/ 1	**-0.10** .2117/ 1	**-0.07** .3392/ 1	**-0.17** .0307/ .9765	**0.02** .7863/ 1
$V_{14\|D1:3}$	**0.27** .0004/ .0588	**0.13** .0990/ 1	**-0.05** .5117/ 1	**0.13** .0799/ 1	**0.04** .6556/ 1	**0.25** .0015/ .1574	**0.18** .0166/ .9013
$V_{15\|D1:3}$	**-0.26** .0006/ .0832	**-0.10** .1848/ 1	**0.00** .9876/ 1	**0.06** .4739/ 1	**-0.10** .1895/ 1	**-0.11** .1445/ 1	**-0.03** .7117/ 1
$V_{16\|D1:3}$	**0.09** .2300/ 1	**0.00** .9566/ 1	**-0.05** .5579/ 1	**-0.20** .0080/ .6390	**-0.03** .7106/ 1	**0.66** .0001/ .0001	**0.02** .8318/ 1
$V_{17\|D1:3}$	**0.10** .1846/ 1	**-0.04** .6037/ 1	**-0.03** .6946/ 1	**0.16** .0353/ .9899	**-0.02** .7869/ 1	**-0.14** .0781/ 1	**0.32** .0001/ .0039

Table S9. Cross-loadings of the imaging side for the 2:1 and 3:1 participant-to-input ratios in the discovery dataset. P-values were obtained using 10000 permutations corrected over 119 tests. Colored cells indicate stability on both the clinical and imaging side. Light green and yellow cells further indicate the CVs highlighted in the main text, which replicate across samples for both clinical and imaging loadings in green or only clinical loadings in purple.

	$u_{1\|D1:2}$	$u_{2\|D1:2}$	$u_{3\|D1:2}$	$u_{4\|D1:2}$	$u_{5\|D1:2}$	$u_{6\|D1:2}$	$u_{7\|D1:2}$
$u_{1\|D1:3}$	**-0.03** .6967/ 1	**0.00** .9696/ 1	**0.84** .0001/ .0001	**0.00** .9961/ 1	**-0.02** .7699/ 1	**0.01** .8612/ 1	**0.00** .9950/ 1
$u_{2\|D1:3}$	**-0.03** .7331/ 1	**-0.78** .0001/ .001	**-0.06** .4132/ 1	**0.05** .5606/ 1	**0.01** .8617/ 1	**0.02** .8360/ 1	**0.01** .8634/ 1
$u_{3\|D1:3}$	**-0.10** .1899/ 1	**-0.08** .3390/ 1	**0.08** .3278/ 1	**0.04** .5969/ 1	**0.02** .7714/ 1	**0.06** .4353/ 1	**-0.05** .5110/ 1
$u_{4\|D1:3}$	**-0.01** .8798/ 1	**0.04** .5746/ 1	**0.04** .6024/ 1	**-0.11** .1472/ 1	**0.00** .9962/ 1	**-0.01** .8780/ 1	**0.00** .9890/ 1
$u_{5\|D1:3}$	**-0.08** .2787/ 1	**-0.02** .8265/ 1	**-0.09** .2657/ 1	**-0.02** .8213/ 1	**0.00** .9620/ 1	**-0.09** .2712/ 1	**-0.04** .5696/ 1
$u_{6\|D1:3}$	**0.00** .9601/ 1	**0.06** .4840/ 1	**-0.07** .3889/ 1	**0.00** .9805/ 1	**0.76** .0001/ .0001	**0.03** .7050/ 1	**0.02** .7680/ 1
$u_{7\|D1:3}$	**-0.32** .0001/ .0043	**0.16** .0370/ .9919	**-0.11** .1438/ 1	**-0.33** .0001/ .0026	**-0.04** .6225/ 1	**-0.04** .6118/ 1	**0.11** .1727/ 1
$u_{8\|D1:3}$	**-0.14** .0710/ 1	**0.03** .7158/ 1	**0.02** .7532/ 1	**-0.15** .0560/ .9994	**0.02** .7987/ 1	**0.00** .9729/ 1	**0.05** .4922/ 1
$u_{9\|D1:3}$	**0.05** .5572/ 1	**-0.04** .6323/ 1	**0.08** .3316/ 1	**-0.12** .1190/ 1	**-0.03** .6945/ 1	**0.00** .9479/ 1	**0.09** .2587/ 1
$u_{10\|D1:3}$	**0.05** .5568/ 1	**0.11** .1481/ 1	**-0.04** .5955/ 1	**0.04** .5832/ 1	**-0.10** .2173/ 1	**0.01** .8630/ 1	**-0.72** .0001/ .0001
$u_{11\|D1:3}$	**-0.09** .2550/ 1	**0.03** .7203/ 1	**-0.08** .3106/ 1	**0.26** .0007/ .0732	**-0.09** .2752/ 1	**0.03** .6987/ 1	**0.06** .4335/ 1
$u_{12\|D1:3}$	**0.05** .5339/ 1	**0.06** .4764/ 1	**-0.03** .7191/ 1	**0.09** .2175/ 1	**-0.09** .2395/ 1	**0.01** .8596/ 1	**-0.01** .9290/ 1
$u_{13\|D1:3}$	**0.02** .7561/ 1	**0.06** .4224/ 1	**0.05** .5379/ 1	**-0.02** .7635/ 1	**-0.12** .1296/ 1	**-0.13** .0882/ 1	**0.00** .9686/ 1
$u_{14\|D1:3}$	**0.17** .0260/ .9573	**0.13** .0933/ 1	**-0.05** .5239/ 1	**0.09** .2359/ 1	**-0.01** .8904/ 1	**0.15** .0541/ .9994	**0.17** .0278/ .9707
$u_{15\|D1:3}$	**−0.01** .3525/ 1	**−0.11** .1131/ 1	**0.07** .8315/ 1	**−0.03** .3253/ 1	**−0.18** .1103/ 1	**−0.02** .1111/ 1	**−0.03** .7405/ 1
$u_{16\|D1:3}$	**-0.07** .3569/ 1	**-0.12** .8726/ 1	**0.02** .8879/ 1	**0.08** .1224/ 1	**-0.13** .9388/ 1	**-0.12** .0001/ .0002	**-0.02** .9635/ 1
$u_{17\|D1:3}$	**0.05** .5444/ 1	**-0.02** .8410/ 1	**-0.04** .6017/ 1	**0.11** .1471/ 1	**-0.03** .6837/ 1	**-0.02** .7876/ 1	**0.27** .0004/ .0432

Table S10. Cross-loadings of the clinical side for the 2:1 and 4:1 participant-to-input ratios in the discovery dataset. P-values were obtained using 10000 permutations corrected over 28 tests. Colored cells indicate stability on both the clinical and imaging side. Light green and yellow cells further indicate the CVs highlighted in the main text, which replicate across samples for both clinical and imaging loadings in green or only clinical loadings in purple.

	$v_{1\|D1:2}$	$v_{2\|D1:2}$	$v_{3\|D1:2}$	$v_{4\|D1:2}$	$v_{5\|D1:2}$	$v_{6\|D1:2}$	$v_{7\|D1:2}$
$v_{1\|D1:4}$	**0.01** .8893/ 1	**-0.04** .6607/ 1	**-0.96** .0001/ .0001	**0.00** .9875/ 1	**-0.03** .6958/ 1	**0.02** .8301/ 1	**0.01** .8652/ 1
$v_{2\|D1:4}$	**0.10** .1846/ .9971	**-0.22** .0041/ .1098	**0.08** .3089/ 1	**-0.15** .0575/ .8025	**-0.91** .0001/ .0001	**-0.02** .7831/ 1	**-0.01** .9425/ 1
$v_{3\|D1:4}$	**-0.03** .7352/ 1	**0.02** .8375/ 1	**-0.07** .3506/ 1	**0.15** .0538/ .7962	**-0.05** .4841/ 1	**0.02** .8220/ 1	**0.00** .9868/ 1
$v_{4\|D1:4}$	**0.53** .0001/ .0001	**-0.31** .0001/ .0007	**0.17** .0287/ .5762	**0.43** .0001/ .0001	**0.10** .2174/ .9988	**0.04** .5953/ 1	**-0.03** .7471/ 1

Table S11. Cross-loadings of the imaging side for the 2:1 and 4:1 participant-to-input ratios in the discovery dataset. P-values were obtained using 10000 permutations corrected over 28 tests. Colored cells indicate stability on both the clinical and imaging side. Light green and yellow cells further indicate the CVs highlighted in the main text, which replicate across samples for both clinical and imaging loadings in green or only clinical loadings in purple.

	$u_{1\|D1:2}$	$u_{2\|D1:2}$	$u_{3\|D1:2}$	$u_{4\|D1:2}$	$u_{5\|D1:2}$	$u_{6\|D1:2}$	$u_{7\|D1:2}$
$u_{1\|D1:4}$	**0.04** .6268/ 1	**0.02** .8024/ 1	**-0.82** .0001/ .0001	**0.01** .9189/ 1	**0.01** .9045/ 1	**0.00** .9555/ 1	**-0.03** .7336/ 1
$u_{2\|D1:4}$	**-0.01** .8985/ 1	**-0.10** .1981/ .9986	**0.06** .4313/ 1	**0.00** .9842/ 1	**-0.72** .0001/ .0001	**-0.03** .7387/ 1	**-0.04** .6313/ 1
$u_{3\|D1:4}$	**0.00** .9823/ 1	**-0.03** .7046/ 1	**-0.04** .5693/ 1	**0.09** .2292/ .9994	**0.00** .9590/ 1	**0.00** .9548/ 1	**-0.01** .9038/ 1
$u_{4\|D1:4}$	**0.30** .0004/ .0031	**-0.19** .0191/ .3822	**0.12** .1368/ .9831	**0.33** .0002/ .0009	**0.06** .4530/ 1	**0.02** .8115/ 1	**0.03** .6541/ 1

Canonical correlations in the replication dataset

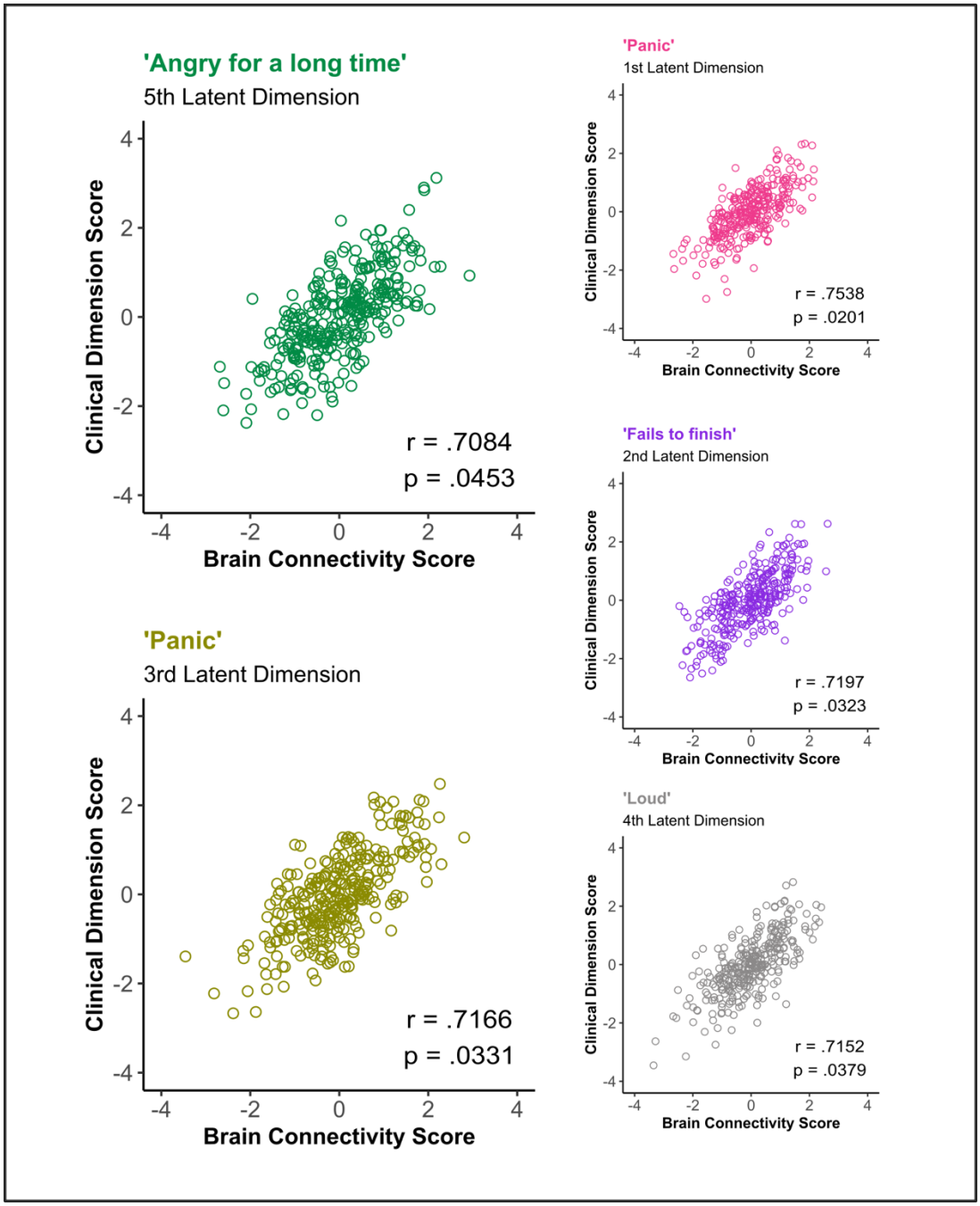


Figure S11. Illustration of the five significant canonical variates in the replication dataset. Scatter plots show the latent dimensions (linear combinations of brain connectivity scores obtained during rsfMRI in the horizontal axis, and linear combinations of clinical scores derived from symptom ratings in the vertical axis). Labels for each latent dimension are derived from the highest loading clinical item. P-values are corrected for multiple testing using the family wise error rate.

Overview of canonical loadings in the replication dataset

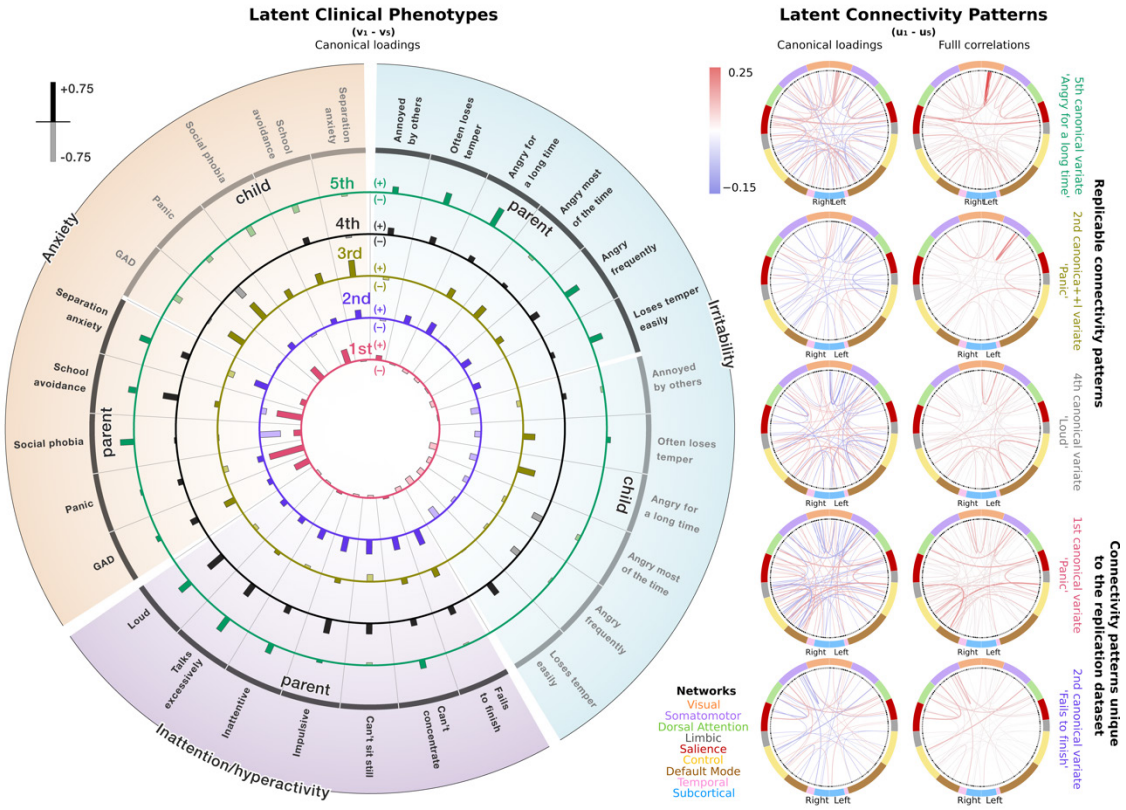
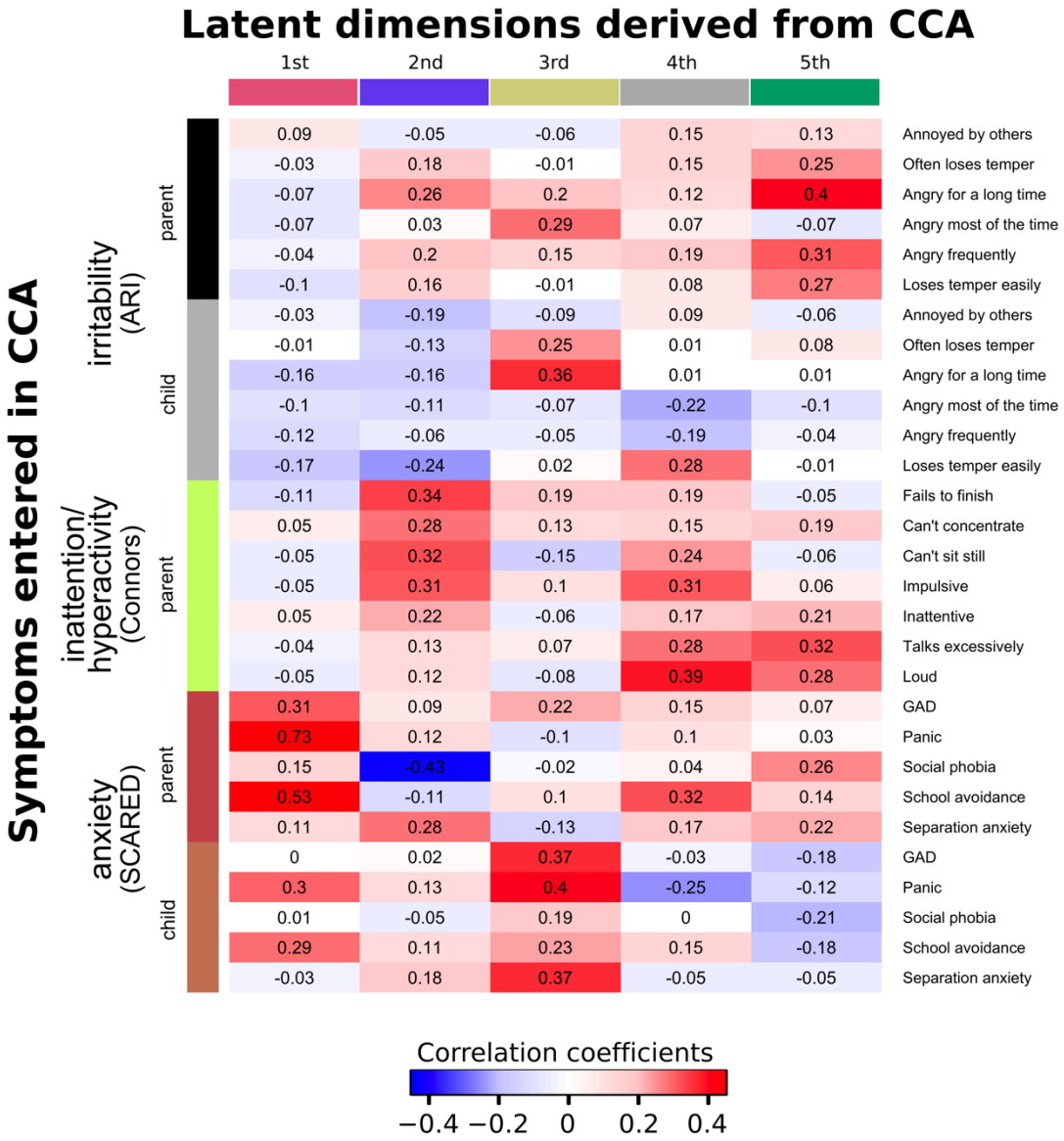


Figure S12. Illustration of the CCA results in the replication sample. The clinical loadings for the five significant canonical variates in the replication dataset are depicted at the left. The connectivity patterns that correlate with the five canonical variates (thresholded at |r| > 0.2) are shown in the middle. Red colored lines represent positive associations between the latent dimension and the edge, whereas blue lines indicate a negative relationship. The intensity of the lines shows the magnitude of the correlation on a scale from -1 to 1, and the thickness reflects the importance of the connection in general (i.e., full correlations). The column to the right shows the directions of the full correlations for that edge, with red indicating positive connectivity and blue negative connectivity.

Clinical loadings in the replication dataset

Table S12. Clinical loadings of the replication dataset



Connectivity loadings in the replication dataset

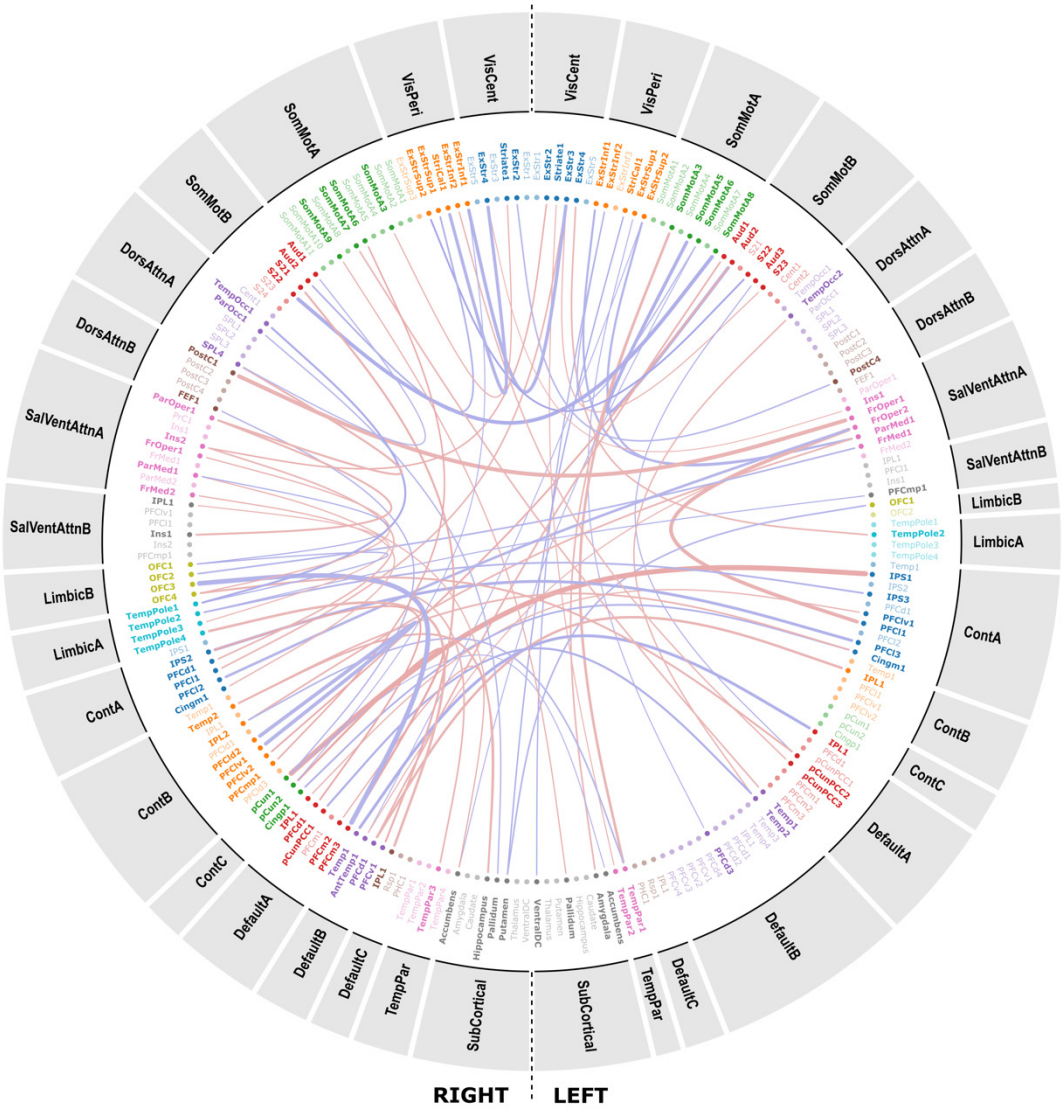


Figure S13. Enlarged depiction of the connectivity pattern associated with the 1st canonical variate in the replication dataset. Edges are thresholded at |r| > 0.2.

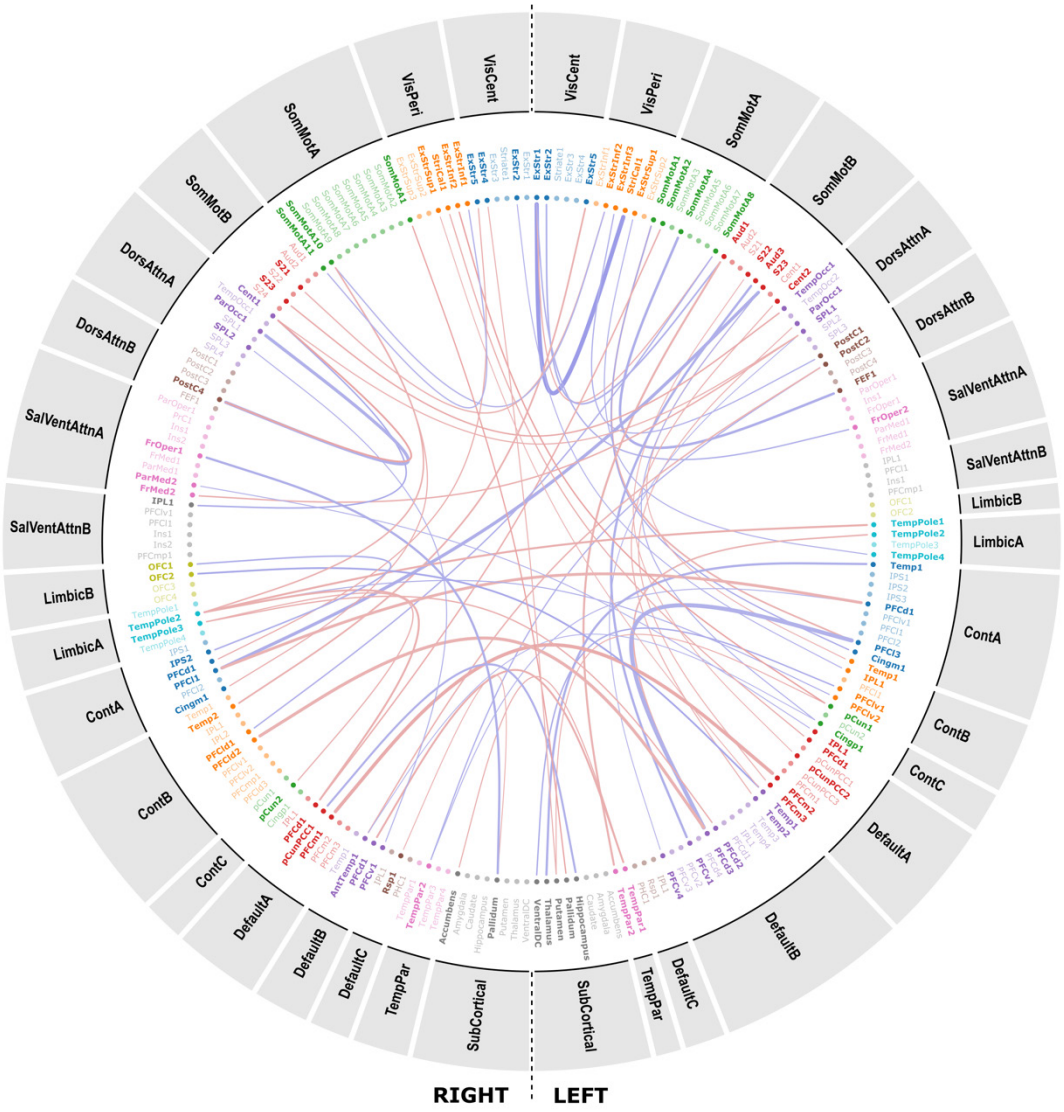


Figure S14. Enlarged depiction of the connectivity pattern associated with the 2nd canonical variate in the replication dataset. Edges are thresholded at |r| > 0.2.

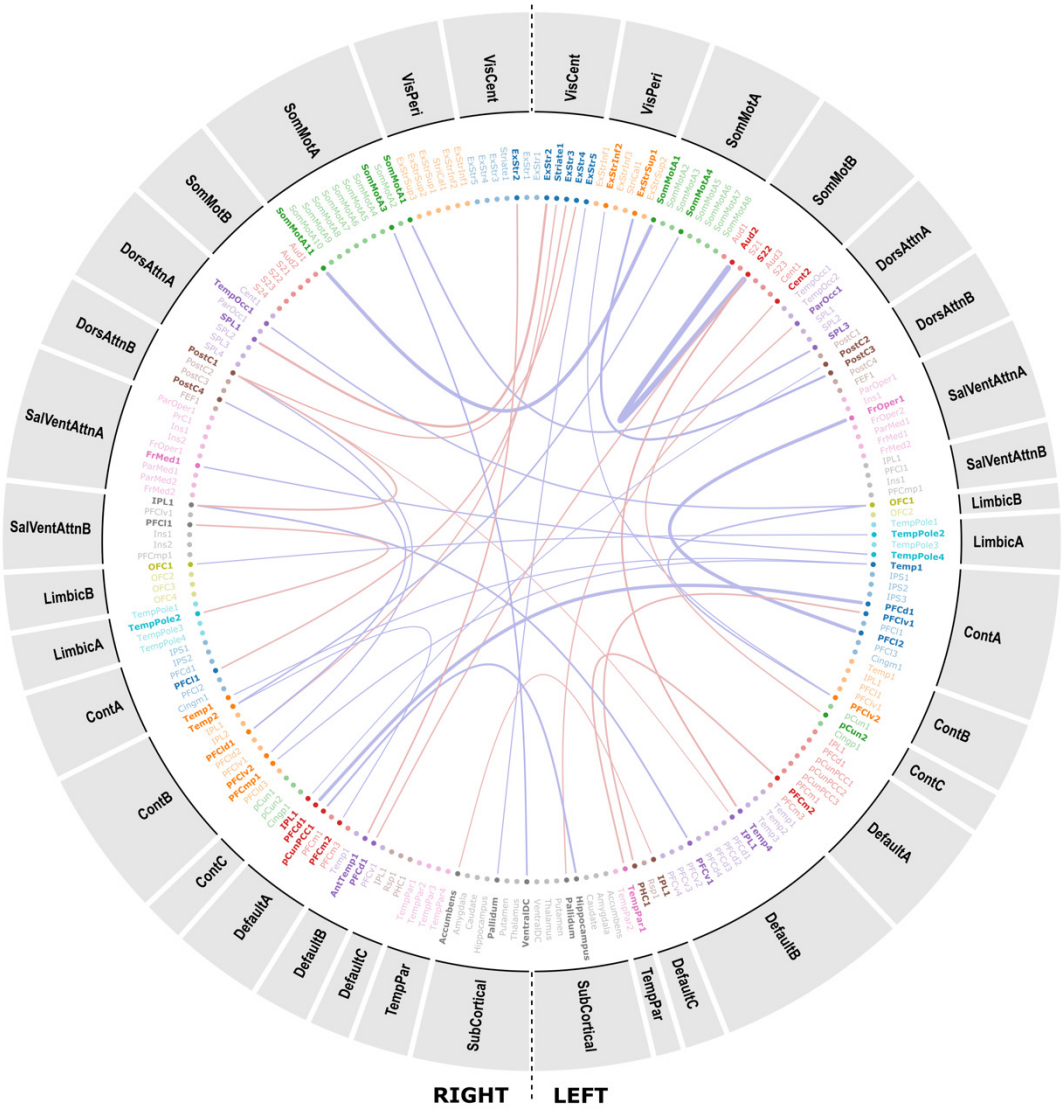


Figure S15. Enlarged depiction of the connectivity pattern associated with the 3rd canonical variate in the replication dataset. Edges are thresholded at |r| > 0.2.

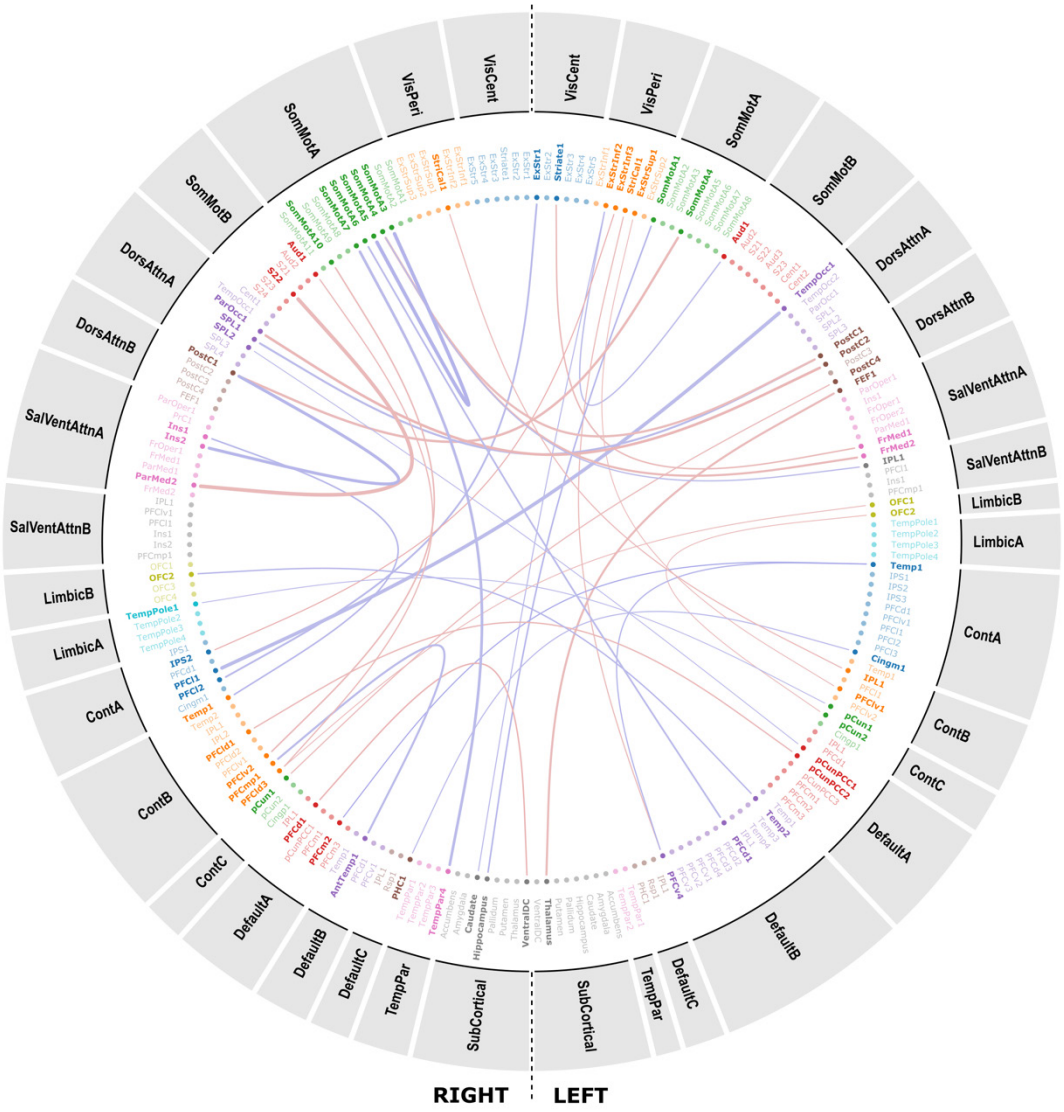


Figure S16. Enlarged depiction of the connectivity pattern associated with the 4th canonical variate in the replication dataset. Edges are thresholded at |r| > 0.2.

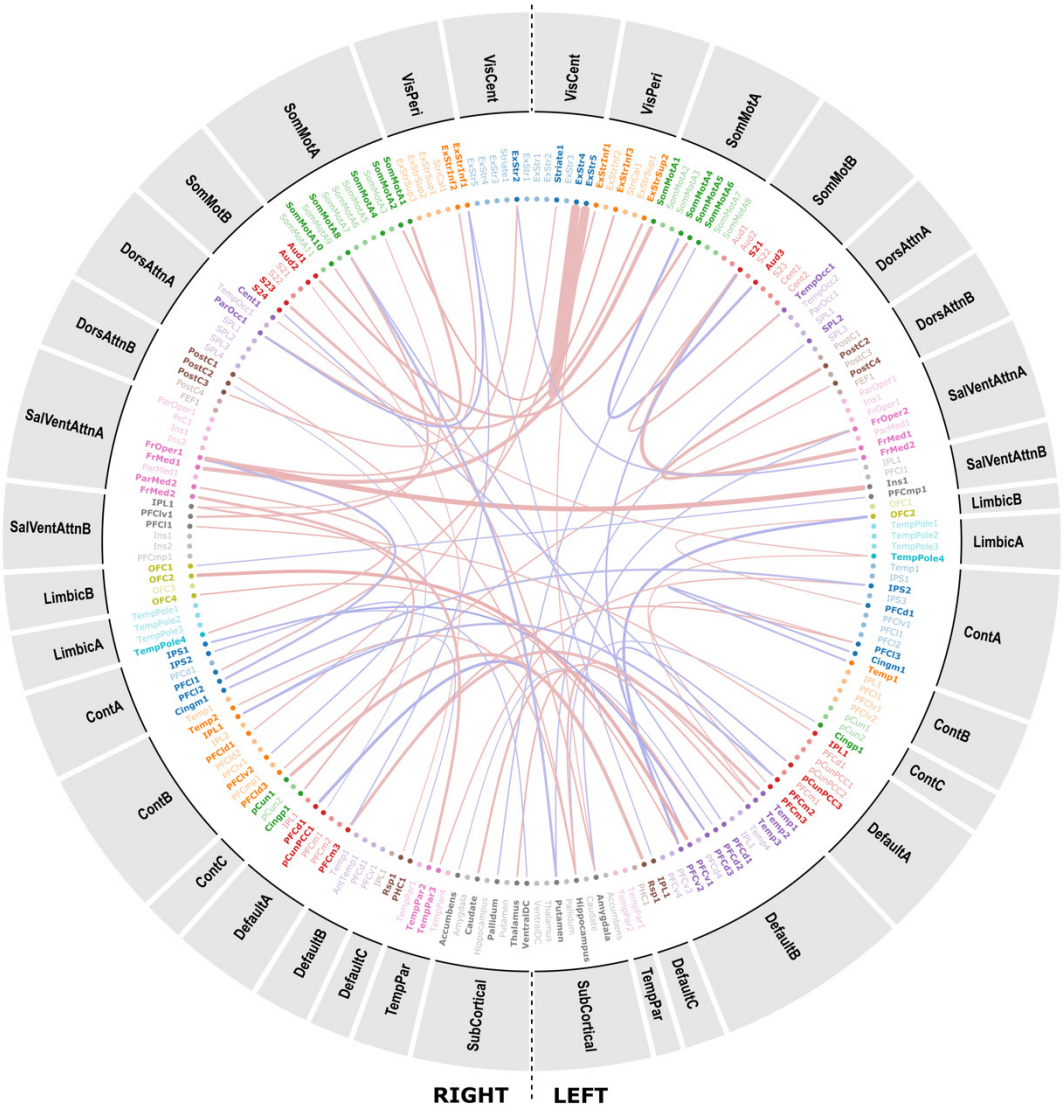


Figure S17. Enlarged depiction of the connectivity pattern associated with the 5th canonical variate in the replication dataset. Edges are thresholded at |r| > 0.2.

Stability of results in the replication dataset

When applying a input-to-participant ratio of 1:3 to select the PCA-components that entered the CCA, CCA yielded 26 significant CVs (CV_1: r = .60, p = 0.0002; CV_2: r = .60, p = 0.0002; CV_3: r =.59 , p = 0.0002 ; CV_4: r = .58, p = 0.0005; CV_5: r = .58, p = 0.0006; CV_6: r = .56, p = 0.0013; CV_7: r = .55, p = 0.0016; CV_8: r = .55, p = 0.0016; CV_9: r = .54, p = 0.0016; CV_{10}: r = .54, p = 0.0018; CV_{11}: r = .53, p = 0.0023; CV_{12}: r = .53, p = 0.0029; CV_{13}: r = .51, p = 0.0037; CV_{14}: r = .51, p = 0.0037; CV_{15}: r = .51, p = 0.0037; CV_{16}: r = .51, p = 0.0037; CV_{17}: r = .50, p = 0.0046; CV_{18}: r = .49, p = 0.0056; CV_{19}: r = .48, p = 0.0066; CV_{20}: r = .47, p = 0.0071; CV_{21}: r = .47, p = 0.0071; CV_{22}: r = .46, p = 0.0071; CV_{23}: r = .46, p = 0.0071; CV_{24}: r = .43, p = 0.0071; CV_{25}: r = .41, p = 0.0112; CV_{26}: r = .41, p = 0.0283). When applying a input-to-participant ratio of 1:4 to the to select the PCA-components that entered the CCA, CCA yielded 22 significant CVs (CV_1: r = .53, p = 0.003; CV_2: r = .48, p = 0.0055; CV_3: r =.48 , p = 0.0061; CV_4: r = .47, p = 0.0063; CV_5: r = .46, p = 0.0070; CV_6: r = .46, p = 0.0072; CV_7: r = .46, p = 0.0084; CV_8: r = .44, p = 0.0096; CV_9: r = .44, p = 0.0107; CV_{10}: r = .43, p = 0.0131; CV_{11}: r = .42, p = 0.0143; CV_{12}: r = .41, p = 0.0143; CV_{13}: r = .41, p = 0.0143; CV_{14}: r = .41, p = 0.0143; CV_{15}: r = .40, p = 0.0143; CV_{16}: r = .40, p = 0.0143; CV_{17}: r = .39, p = 0.0143; CV_{18}: r = .39, p = 0.0143; CV_{19}: r = .39, p = 0.0143; CV_{20}: r = .39, p = 0.0143; CV_{21}: r = .38, p = 0.0315; CV_{22}: r = .37, p = 0.0369).

Table S13. Cross-loadings of the clinical side across participant-to-input ratios in the replication dataset. P-values were obtained using 10000 permutations corrected over 130 tests. Colored cells indicate stability on both the clinical and imaging side. Light green and purple cells further indicate the CVs highlighted in the main text, which replicate across samples for both clinical (and imaging data).

	$V_{1\|R1:2}$	$V_{2\|R1:2}$	$V_{3\|R1:2}$	$V_{4\|R1:2}$	$V_{5\|R1:2}$
$V_{1\|R1:3}$	**0.01** .8635/ 1	**0.00** .9876/ 1	**0.06** .3026/ 1	**0.02** .7327/ 1	**-0.06** .2786/ 1
$V_{2\|R1:3}$	**0.98** .0001/ .0001	**0.07** .2205/ 1	**-0.03** .6440/ 1	**-0.03** .6444/ 1	**-0.01** .8065/ 1
$V_{3\|R1:3}$	**0.02** .7348/ 1	**-0.13** .0168/ .9004	**-0.35** .0001/ .0001	**0.04** .5275/ 1	**-0.06** .2543/ 1
$V_{4\|R1:3}$	**0.03** .5713/ 1	**-0.01** .8867/ 1	**-0.11** .0584/ .9998	**-0.08** .1368/ 1	**-0.05** .3955/ 1
$V_{5\|R1:3}$	**-0.03** .5898/ 1	**-0.05** .4010/ 1	**-0.17** .0038/ .3587	**-0.04** .5025/ 1	**0.04** .4684/ 1
$V_{6\|R1:3}$	**-0.04** .4799/ 1	**0.14** .0133/ .8340	**-0.03** .5591/ 1	**0.04** .4433/ 1	**0.04** .5299/ 1
$V_{7\|R1:3}$	**0.04** .4934/ 1	**-0.88** .0001/ .0001	**-0.18** .0011/ .1617	**0.07** .2235/ 1	**-0.12** .0399/ .9961
$V_{8\|R1:3}$	**0.00** .9607/ 1	**-0.10** .0764/ .9999	**-0.05** .4245/ 1	**-0.07** .2228/ 1	**0.98** .0001/ .0001
$V_{9\|R1:3}$	**0.05** .3473/ 1	**-0.06** .3292/ 1	**-0.01** .8416/ 1	**0.16** .0049/ .4651	**-0.01** .8880/ 1
$V_{10\|R1:3}$	**-0.02** .7030/ 1	**-0.11** .0612/ .9999	**0.08** .1544/ 1	**-0.12** .0424/ .9957	**-0.02** .7350/ 1
$V_{11\|R1:3}$	**0.01** .9152/ 1	**0.10** .0733/ .9998	**-0.37** .0001/ .0001	**0.02** .7162/ 1	**0.03** .6305/ 1
$V_{12\|R1:3}$	**0.03** .7860/ 1	**-0.03** .5929/ 1	**0.03** .9814/ 1	**0.28** .0354/ .9889	**0.06** .4499/ 1
$V_{13\|R1:3}$	**0.01** .7978/ 1	**0.05** .3800/ 1	**-0.03** .5762/ 1	**0.47** .0001/ .0001	**0.06** .3273/ 1
$V_{14\|R1:3}$	**0.02** .7616/ 1	**0.05** .3930/ 1	**-0.10** .0689/ .9998	**-0.20** .0004/ .0543	**0.00** .9936/ 1
$V_{15\|R1:3}$	**0.05** .4238/ 1	**-0.03** .6474/ 1	**0.06** .3205/ 1	**0.48** .0001/ .0001	**0.02** .6687/ 1
$V_{16\|R1:3}$	**-0.03** .5573/ 1	**0.05** .3517/ 1	**0.11** .0442/ .9977	**0.24** .0002/ .0043	**-0.01** .9161/ 1
$V_{17\|R1:3}$	**0.04** .5157/ 1	**0.00** .9367/ 1	**0.06** .2663/ 1	**0.15** .0103/ .7288	**-0.02** .7443/ 1
$V_{18\|R1:3}$	**0.03** .5815/ 1	**0.03** .6292/ 1	**-0.16** .0051/ .4973	**-0.19** .0005/ .0970	**-0.05** .3569/ 1
$V_{19\|R1:3}$	**-0.03** .5776/ 1	**-0.15** .0100/ .7221	**0.10** .0784/ 1	**-0.06** .3211/ 1	**-0.05** .3392/ 1
$V_{20\|R1:3}$	**0.02** .6762/ 1	**-0.11** .0466/ .9993	**0.41** .0001/ .0001	**0.18** .0020/ .2227	**0.03** .5636/ 1
$V_{21\|R1:3}$	**0.06** .3237/ 1	**0.05** .3371/ 1	**0.40** .0001/ .0001	**-0.12** .0273/ .9799	**-0.05** .3984/ 1
$V_{22\|R1:3}$	**-0.06** .2725/ 1	**0.00** .9846/ 1	**-0.03** .5755/ 1	**0.25** .0002/ .0022	**-0.03** .6123/ 1
$V_{23\|R1:3}$	**-0.05** .3791/ 1	**-0.05** .3942/ 1	**0.22** .0002/ .0099	**-0.09** .1283/ 1	**-0.02** .7009/ 1
$V_{24\|R1:3}$	**-0.05** .4150/ 1	**0.01** .9277/ 1	**0.12** .0368/ .9914	**-0.21** .0003/ .0214	**0.00** .9653/ 1
$V_{25\|R1:3}$	**-0.01** .8086/ 1	**0.01** .8567/ 1	**0.15** .0075/ .6626	**0.19** .0005/ .0778	**0.03** .6195/ 1
$V_{26\|R1:3}$	**-0.04** .5030/ 1	**-0.06** .3091/ 1	**0.13** .0228/ .9607	**0.00** .9697/ 1	**0.02** .6673/ 1

Table S14. Cross-loadings of the imaging side across participant-to-input ratios in the replication dataset. P-values were obtained using 10000 permutations corrected over 130 tests. Colored cells indicate stability on both the clinical and imaging side. Light green and purple cells further indicate the CVs highlighted in the main text, which replicate across samples for both clinical (and imaging data).

	$u_{1\|R1:2}$	$u_{2\|R1:2}$	$u_{3\|R1:2}$	$u_{4\|R1:2}$	$u_{5\|R1:2}$
$u_{1\|R1:3}$	**0.00** .9883/ 1	**-0.02** .7085/ 1	**0.02** .6755/ 1	**0.01** .8123/ 1	**-0.03** .5448/ 1
$u_{2\|R1:3}$	**0.80** .0001/ .0001	**-0.01** .8394/ 1	**-0.06** .2706/ 1	**0.01** .8737/ 1	**-0.01** .8993/ 1
$u_{3\|R1:3}$	**0.05** .4032/ 1	**-0.11** .0555/ .9993	**-0.33** .0001/ .0001	**0.04** .4373/ 1	**-0.07** .2260/ 1
$u_{4\|R1:3}$	**0.07** .2231/ 1	**-0.01** .9272/ 1	**-0.08** .1548/ 1	**-0.08** .1520/ 1	**-0.07** .2395/ 1
$u_{5\|R1:3}$	**-0.02** .7231/ 1	**0.01** .8379/ 1	**-0.14** .0130/ .8657	**-0.02** .7790/ 1	**-0.01** .8149/ 1
$u_{6\|R1:3}$	**-0.03** .6228/ 1	**0.15** .0077/ .6207	**-0.04** .4717/ 1	**0.05** .4144/ 1	**0.03** .5519/ 1
$u_{7\|R1:3}$	**0.11** .0657/ .9998	**-0.71** .0001/ .0001	**-0.15** .0111/ .6768	**0.06** .3014/ 1	**-0.13** .0178
$u_{8\|R1:3}$	**0.01** .8956/ 1	**-0.03** .5671/ 1	**-0.02** .7441/ 1	**0.00** .9379/ 1	**0.77** .0001/ .0001
$u_{9\|R1:3}$	**0.06** .3334/ 1	**-0.08** .1641/ 1	**-0.03** .5903/ 1	**0.10** .0861/ .9999	**0.01** .8691/ 1
$u_{10\|R1:3}$	**0.01** .8346/ 1	**-0.07** .1940/ 1	**0.08** .1581/ 1	**-0.09** .0980/ .9999	**0.01** .8348/ 1
$u_{11\|R1:3}$	**0.04** .4774/ 1	**0.11** .0492/ .9991	**-0.31** .0001/ .0001	**0.01** .8381/ 1	**0.05** .4071/ 1
$u_{12\|R1:3}$	**0.02** .6965/ 1	**-0.04** .5365/ 1	**-0.01** .9054/ 1	**0.19** .0008/ .0860	**0.06** .3027/ 1
$u_{13\|R1:3}$	**0.00** .9988/ 1	**0.01** .7940/ 1	**0.01** .8640/ 1	**0.36** .0001/ .0001	**0.08** .1690/ 1
$u_{14\|R1:3}$	**0.04** .5416/ 1	**0.04** .4890/ 1	**-0.06** .2614/ 1	**-0.08** .1694/ 1	**0.02** .7417/ 1
$u_{15\|R1:3}$	**0.05** .4257/ 1	**-0.03** .6431/ 1	**-0.02** .7942/ 1	**0.35** .0001/ .0001	**0.03** .6023/ 1
$u_{16\|R1:3}$	**-0.03** .5588/ 1	**0.07** .2307/ 1	**0.08** .1419/ 1	**0.11** .0472/ .9988	**-0.01** .8038/ 1
$u_{17\|R1:3}$	**0.02** .7197/ 1	**0.00** .9409/ 1	**0.05** .3817/ 1	**0.11** .0595/ .9998	**-0.04** .4757/ 1
$u_{18\|R1:3}$	**0.02** .6911/ 1	**0.00** .9825/ 1	**-0.11** .0578/ .9997	**-0.20** .0005/ .0563	**-0.08** .1860/ 1
$u_{19\|R1:3}$	**-0.03** .6335/ 1	**-0.10** .0695/ .9998	**0.13** .0215/ .9468	**0.03** .6093/ 1	**-0.08** .1817/ 1
$u_{20\|R1:3}$	**0.04** .5141/ 1	**-0.08** .1830/ 1	**0.32** .0001/ .0001	**0.09** .0978/ .9999	**0.05** .3660/ 1
$u_{21\|R1:3}$	**0.04** .4979/ 1	**0.09** .1350/ 1	**0.30** .0001/ .0002	**-0.13** .0164/ .9190	**-0.06** .3173/ 1
$u_{22\|R1:3}$	**-0.07** .2232/ 1	**0.00** .9827/ 1	**-0.05** .3348/ 1	**0.23** .0001/ .0079	**-0.02** .6582/ 1
$u_{23\|R1:3}$	**-0.08** .1454/ 1	**-0.02** .7689/ 1	**0.17** .0036/ .3539	**-0.01** .8201/ 1	**-0.05** .3722/ 1
$u_{24\|R1:3}$	**-0.08** .1864/ 1	**0.03** .6013/ 1	**0.09** .0941/ .9999	**-0.15** .0053/ .5622	**-0.04** .4904/ 1
$u_{25\|R1:3}$	**0.01** .8529/ 1	**0.04** .5154/ 1	**0.13** .0276/ .9682	**0.10** .0732/ .9999	**0.04** .4853/ 1
$u_{26\|R1:3}$	**-0.02** .7546/ 1	**-0.07** .2378/ 1	**0.07** .2166/ 1	**0.01** .8215/ 1	**0.05** .4131/ 1

Table S15. Cross-loadings of the clinical side for the 2:1 and 4:1 participant-to-input ratios in the replication dataset. P-values were obtained using 10000 permutations corrected over 110 tests. Colored cells indicate stability on both the clinical and imaging side. Light green and yellow cells further indicate the CVs highlighted in the main text, which replicate across samples for both clinical and imaging loadings in green or only clinical loadings in purple.

	$V_{1\|R1:2}$	$V_{2\|R1:2}$	$V_{3\|R1:2}$	$V_{4\|R1:2}$	$V_{5\|R1:2}$
$V_{1\|R1:4}$	**0.01** .9117/ 1	**0.01** .8510/ 1	**0.05** .4299/ 1	**0.09** .1061	**0.04** .4425/ 1
$V_{2\|R1:4}$	**0.01** .8315/ 1	**0.03** .6251/ 1	**0.13** .0188/ .8613	**0.05** .3975/ 1	**-0.03** .6395/ 1
$V_{3\|R1:4}$	**-0.02** .7477/ 1	**0.11** .0530/ .9990	**0.11** .0637/ .9993	**0.27** .0001/ .0003	**0.12** .0324/ .9721
$V_{4\|R1:4}$	**0.02** .7063/ 1	**0.05** .3415/ 1	**0.17** .0038/ .3239	**0.05** .3655/ 1	**-0.93** .0001/ .0001
$V_{5\|R1:4}$	**-0.13** .0194/ .8969	**-0.16** .0049/ .3772	**0.05** .3828/ 1	**-0.08** .1606/ 1	**-0.02** .6594/ 1
$V_{6\|R1:4}$	**-0.20** .0010/ .0393	**0.23** .0001/ .0001	**-0.15** .0109/ .6784	**-0.14** .0114/ .7239	**0.03** .6518/ 1
$V_{7\|R1:4}$	**-0.19** .0009/ .0879	**-0.05** .3374/ 1	**-0.01** .8557/ 1	**-0.18** .0016/ .1797	**-0.04** .5308/ 1
$V_{8\|R1:4}$	**0.06** .2620/ 1	**0.34** .0001/ .0001	**0.12** .0307/ .9684	**-0.02** .7375/ 1	**-0.08** .1422/ 1
$V_{9\|R1:4}$	**-0.06** .3147/ 1	**0.03** .5911/ 1	**0.18** .0022/ .1702	**-0.08** .1647/ 1	**-0.02** .7633/ 1
$V_{10\|R1:4}$	**-0.07** .2472/ 1	**0.03** .5760/ 1	**-0.18** .0011/ .1183	**0.31** .0001/ .0001	**-0.07** .2123/ 1
$V_{11\|R1:4}$	**0.09** .1006/ .9999	**0.62** .0001/ .0001	**0.15** .0068/ .5678	**-0.33** .0001/ .0001	**0.09** .1176/ 1
$V_{12\|R1:4}$	**-0.10** .0809/ .9999	**-0.01** .7912/ 1	**0.12** .0294/ .9682	**-0.05** .4131/ 1	**-0.03** .6408/ 1
$V_{13\|R1:4}$	**0.55** .0001/ .0001	**-0.01** .8629/ 1	**0.06** .3252/ 1	**-0.19** .0005/ .0930	**-0.08** .1608/ 1
$V_{14\|R1:4}$	**-0.12** .0387/ .9885	**0.24** .0001/ .0001	**0.17** .0029/ .2786	**0.53** .0001/ .0001	**0.03** .6400/ 1
$V_{15\|R1:4}$	**0.70** .0001/ .0001	**0.04** .4857/ 1	**-0.14** .0127/ .7221	**0.13** .0161/ .8683	**0.00** .9721/ 1
$V_{16\|R1:4}$	**-0.04** .4871/ 1	**-0.27** .0001/ .0002	**-0.05** .3766/ 1	**-0.10** .0712/ .9996	**0.01** .8212/ 1
$V_{17\|R1:4}$	**0.11** .0552/ .9909	**-0.05** .3432/ 1	**0.16** .0057/ .4407	**0.07** .2127/ 1	**0.07** .2460/ 1
$V_{18\|R1:4}$	**-0.11** .0563/ .9990	**0.24** .0001/ .0011	**-0.45** .0001/ .0001	**-0.06** .2892/ 1	**-0.10** .0912/ .9999
$V_{19\|R1:4}$	**-0.06** .3340/ 1	**0.07** .2061/ 1	**0.14** .0168/ .8434	**0.20** .0005/ .0326	**-0.05** .3686/ 1
$V_{20\|R1:4}$	**0.04** .5098/ 1	**-0.03** .5677/ 1	**0.14** .0178/ .8585	**-0.18** .0016/ .1292	**0.01** .8000/ 1
$V_{21\|R1:4}$	**0.11** .0433/ .9936	**-0.09** .1195/ 1	**-0.21** .0002/ .0161	**0.19** .0006/ .0972	**-0.02** .6911/ 1
$V_{22\|R1:4}$	**-0.13** .0295/ .9517	**-0.10** .0744/ .9997	**0.14** .0151/ .8318	**-0.11** .0532/ .9984	**-0.02** .7467/ 1

Table S16. Cross-loadings of the imaging side for the 2:1 and 4:1 participant-to-input ratios in the replication dataset. P-values were obtained using 10000 permutations corrected over 110 tests. Colored cells indicate stability on both the clinical and imaging side. Light green and yellow cells further indicate the CVs highlighted in the main text, which replicate across samples for both clinical and imaging loadings in green or only clinical loadings in purple.

	$u_{1\|R1:2}$	$u_{2\|R1:2}$	$u_{3\|R1:2}$	$u_{4\|R1:2}$	$u_{5\|R1:2}$
$u_{1\|R1:4}$	**0.00** .9432/ 1	**-0.01** .8768/ 1	**-0.01** .8767/ 1	**0.05** .3687/ 1	**0.04** .4982/ 1
$u_{2\|R1:4}$	**0.03** .6488/ 1	**-0.02** .6995/ 1	**0.12** .0314/ .9664	**0.01** .8954/ 1	**0.02** .6993/ 1
$u_{3\|R1:4}$	**-0.03** .5514/ 1	**0.13** .0194/ .9020	**0.04** .4703/ 1	**0.11** .0514/ .9968	**0.11** .0536/ .9986
$u_{4\|R1:4}$	**-0.01** .8119/ 1	**-0.01** .9283/ 1	**0.04** .4527/ 1	**0.00** .9547/ 1	**-0.67** .0001/ .0001
$u_{5\|R1:4}$	**-0.10** .0607/ .9997	**-0.13** .0222/ .9157	**0.02** .7312/ 1	**-0.07** .2354/ 1	**-0.03** .6351/ 1
$u_{6\|R1:4}$	**-0.11** .0646/ .9995	**0.22** .0002/ .0165	**-0.02** .6743/ 1	**-0.08** .1410/ 1	**0.02** .7181/ 1
$u_{7\|R1:4}$	**-0.16** .0066/ .4823	**-0.04** .5102/ 1	**0.03** .6620/ 1	**-0.07** .2115/ 1	**-0.07** .2479/ 1
$u_{8\|R1:4}$	**-0.02** .7731/ 1	**0.29** .0001/ .0001	**0.06** .2762/ 1	**0.00** .9554/ 1	**-0.06** .2568/ 1
$u_{9\|R1:4}$	**-0.06** .2750/ 1	**0.06** .3060/ 1	**0.08** .1539/ 1	**-0.03** .5760/ 1	**-0.02** .7779/ 1
$u_{10\|R1:4}$	**-0.03** .6021/ 1	**-0.01** .9294/ 1	**-0.11** .0577/ .9987	**0.22** .0002/ .0098	**-0.05** .3725/ 1
$u_{11\|R1:4}$	**0.06** .3282/ 1	**0.43** .0001/ .0001	**0.11** .0530/ .9976	**-0.18** .0009/ .1420	**0.10** .0650/ .9996
$u_{12\|R1:4}$	**-0.14** .0162/ .8323	**-0.03** .6267/ 1	**0.07** .2049/ 1	**-0.06** .2618/ 1	**-0.08** .1803/ 1
$u_{13\|R1:4}$	**0.44** .0001/ .0001	**-0.04** .4772/ 1	**0.02** .6733/ 1	**-0.07** .2357/ 1	**-0.03** .6587/ 1
$u_{14\|R1:4}$	**-0.12** .0377/ .9880	**0.17** .0027/ .2617	**0.12** .0283/ .9614	**0.35** .0001/ .0001	**0.03** .6418/ 1
$u_{15\|R1:4}$	**0.48** .0001/ .0001	**0.00** .9320/ 1	**-0.11** .0632/ .9995	**0.06** .2671/ 1	**-0.02** .6695/ 1
$u_{16\|R1:4}$	**-0.08** .1617/ 1	**-0.23** .0002/ .0086	**0.04** .5338/ 1	**-0.08** .1856/ 1	**-0.03** .6508/ 1
$u_{17\|R1:4}$	**0.10** .0873/ 1	**-0.05** .3951/ 1	**0.11** .0511/ .9967	**-0.03** .6354/ 1	**0.07** .2491/ 1
$u_{18\|R1:4}$	**-0.12** .0287/ .9652	**0.12** .0380/ .9782	**-0.17** .0029/ .2838	**-0.01** .8738/ 1	**-0.09** .1088/ 1
$u_{19\|R1:4}$	**-0.03** .5489/ 1	**0.03** .6426/ 1	**0.13** .0246/ .9458	**0.11** .0622/ .9995	**-0.03** .6596/ 1
$u_{20\|R1:4}$	**0.06** .3300/ 1	**-0.04** .4704/ 1	**0.08** .1493/ 1	**-0.08** .1784/ 1	**0.02** .7516/ 1
$u_{21\|R1:4}$	**0.12** .0385/ .9910	**-0.09** .1186/ 1	**-0.06** .2640/ 1	**0.11** .0593/ .9995	**-0.03** .6435/ 1
$u_{22\|R1:4}$	**-0.09** .1217/ 1	**-0.08** .1505/ 1	**0.05** .3453/ 1	**-0.01** .8004/ 1	**-0.03** .5888/ 1

Replicability across cohorts

Table S17. Correlations between the significant latent clinical variables ($v_{1\text{-}5}$) from the replication dataset ($_R$) and the clinical data from the replication cohort multiplied by the canonical weights of the significant clinical variables ($b_{1\text{-}7}$) from the discovery dataset ($_D$). In each cell the first value represents the correlation coefficient, the second value gives the uncorrected p-value, and the third value shows the p_{FWER}-value. Light green cells indicate robust cross-correlations for both clinical and imaging CVs, whereas light purple cells indicate robust cross-correlations for the clinical CVs only.

	$v_{1\|R}$	$v_{2\|R}$	$v_{3\|R}$	$v_{4\|R}$	$v_{5\|R}$
$X_R \times b_{1\|D}$	**0.08** .1360/ .9944	**0.11** .0528/ .8430	**−0.13** .0256/ .5831	**−0.26** .0001/ .0001	**0.33** .0001/ .0001
$X_R \times b_{2\|D}$	**0.08** .1516/ .9954	**−0.22** .0001/ .0024	**0.26** .0001/ .0001	**−0.04** .4633/ 1	**0.25** .0001/ .0004
$X_R \times b_{3\|D}$	**−0.20** .0008/ .0093	**0.13** .0214/ .5061	**−0.08** .1417/ .9941	**0.32** .0001/ .0001	**-0.20** .0006/ .0175
$X_R \times b_{4\|D}$	**0.28** .0001/ .0001	**0.01** .8646/ 1	**−0.15** .0089/ .2613	**0.19** .0006/ .0229	**−0.10** .0861/ .9559
$X_R \times b_{5\|D}$	**−0.14** .0130/ .3477	**−0.07** .2114/ .9997	**0.00** .9601/ 1	**−0.02** .7398/ 1	**0.22** .0004/ .0022
$X_R \times b_{6\|D}$	**−0.04** .5347/ 1	**−0.06** .2766/ 1	**0.09** .0965/ .9672	**0.10** .0830/ .9390	**0.16** .0045/ .1237
$X_R \times b_{7\|D}$	**0.51** .0001/ .0001	**−0.29** .0001/ .0001	**0.44** .0001/ .0001	**−0.26** .0001/ .0001	**0.02** .7305/ 1

Table S18. Correlations between the significant latent resting-state connectivity variables ($u_{1\text{-}5}$) from the replication dataset ($_R$) and the resting-state data from the replication cohort multiplied by the canonical weights of the significant resting-state connectivity variables ($a_{1\text{-}7}$) from the discovery dataset ($_D$). In each cell the first value represents the correlation coefficient, the second value gives the uncorrected p-value, and the third value shows the p_{FWER}-value. Light green cells indicate robust cross-correlations for both clinical and imaging CVs.

	$u_{1\|R}$	$u_{2\|R}$	$u_{3\|R}$	$u_{4\|R}$	$u_{5\|R}$
$Y_R \times a_{1\|D}$	**0.03** .6312/ 1	**0.03** .6257/ 1	**0.03** .9646/ 1	**−0.04** .4690/ 1	**−0.01** .8429/ 1
$Y_R \times a_{2\|D}$	**−0.02** .7744/ 1	**0.04** .4609/ 1	**0.00** .9819/ 1	**−0.01** .8981/ 1	**−0.03** .6355/ 1
$Y_R \times a_{3\|D}$	**0.02** .6792/ 1	**0.06** .2748/ 1	**0.01** .7990/ 1	**−0.00** .9845/ 1	**0.12** .0375/ .7425
$Y_R \times a_{4\|D}$	**−0.11** .0650/ .8938	**0.07** .2515/ .9998	**−0.07** .2214/ .9998	**0.06** .2518/ .9998	**0.02** .7046/ 1
$Y_R \times a_{5\|D}$	**−0.01** .8061/ 1	**0.08** .1827/ .9989	**0.01** .8101/ .9977	**−0.08** .1655/ .9977	**−0.03** .5406/ 1
$Y_R \times a_{6\|D}$	**−0.01** .9136/ 1	**0.07** .1957/ .9991	**0.07** .1965/ 1	**0.00** .9723/ 1	**0.02** .7862/ 1
$Y_R \times a_{7\|D}$	**−0.08** .1512/ .9962	**0.08** .1499/ .9967	**0.10** .0696/ 1	**−0.00** .9897/ 1	**0.00** .9669/ 1

Table S19. Correlations between the significant latent clinical variables ($v_{1\text{-}7}$) from the discovery dataset ($_D$) and the clinical data from the discovery cohort multiplied by the canonical weights of the significant clinical variables ($b_{1\text{-}5}$) from the replication dataset ($_R$). In each cell the first value represents the correlation coefficient, the second value gives the uncorrected p-value, and the third value shows the p_{FWER}-value. Light green cells indicate robust cross-correlations for both clinical and imaging CVs, whereas light purple cells indicate robust cross-correlations for the clinical CVs only.

	$X_D \times b_{1\|R}$	$X_D \times b_{2\|R}$	$X_D \times b_{3\|R}$	$X_D \times b_{4\|R}$	$X_D \times b_{5\|R}$
$v_{1\|D}$	**0.08** .2999/ 1	**0.02** .7689/ 1	**0.12** .1332/ .9922	**0.15** .0625/ .8859	**-0.16** .0430/ .8030
$v_{2\|D}$	**0.06** .4087/ 1	**0.08** .3254/ 1	**0.13** .0934/ .9693	**−0.01** .9068/ 1	**0.45** .0001/ .0001
$v_{3\|D}$	**−0.15** .0531/ .8200	**−0.10** .1811/ .9988	**0.28** .0003/ .0106	**-0.24** .0019/ .0698	**-0.19** .0141/ .3797
$v_{4\|D}$	**-0.17** .0232/ .5884	**0.02** .7919/ 1	**−0.04** .5782/ 1	**0.27** .0009/ .0190	**-0.10** .1877/ .9993
$v_{5\|D}$	**0.01** .8800/ 1	**−0.16** .0497/ .8045	**0.17** .0257/ .6148	**0.06** .4768/ 1	**0.12** .1265/ .9905
$v_{6\|D}$	**-0.09** .2344/ .9999	**-0.08** .2919/ 1	**0.22** .0033/ .1240	**−0.06** .4601/ 1	**-0.16** .0362/ .7287
$v_{7\|D}$	**0.45** .0001/ .0001	**0.19** .0205/ .4473	**0.31** .0001/ .0017	**−0.16** .0385/ .7313	**-0.02** .7869/ 1

Table S20. Correlations between the significant latent resting-state connectivity variables ($u_{1\text{-}7}$) from the discovery dataset ($_D$) and the resting-state data from the discovery cohort multiplied by the canonical weights of the significant resting-state connectivity variables ($a_{1\text{-}5}$) from the replication dataset ($_R$). In each cell the first value represents the correlation coefficient, the second value gives the uncorrected p-value, and the third value shows the p_{FWER}-value. Light green cell color indicates robust cross-correlations for both clinical and imaging CVs.

	$Y_D \times a_{1\|R}$	$Y_D \times a_{2\|R}$	$Y_D \times a_{3\|R}$	$Y_D \times a_{4\|R}$	$Y_D \times a_{5\|R}$
$u_{1\|D}$	**0.06** .4236/ 1	**0.10** .2204/ .9996	**0.09** .2748/ 1	**0.09** .2655/ 1	**−0.03** .6609/ 1
$u_{2\|D}$	**-0.07** .3673/ 1	**-0.07** .3673/ 1	**−0.05** .5121/ 1	**−0.04** .6400/ 1	**0.00** .9935/ 1
$u_{3\|D}$	**−0.04** .6009/ 1	**0.16** .0375/ .7224	**-0.03** .7488/ 1	**-0.11** .1665/ .9982	**0.32** .0001/ .0007
$u_{4\|D}$	**0.09** .2318/ .9997	**-0.08** .3321/ 1	**−0.19** .0212/ .5444	**0.19** .0108/ .3494	**-0.12** .1324/ .9945
$u_{5\|D}$	**0.04** .6039/ 1	**0.05** .5060/ 1	**0.00** .9506/ 1	**0.07** .5908/ 1	**−0.10** .1796/ .9987
$u_{6\|D}$	**-0.06** .4066/ 1	**-0.12** .1140/ .9860	**0.06** .4728/ 1	**−0.04** .5908/ 1	**-0.04** .6315/ 1
$u_{7\|D}$	**0.08** .3188/ 1	**-0.13** .0904/ .9612	**0.19** .0136/ .4120	**−0.03** .7530/ 1	**-0.02** .7905/ 1

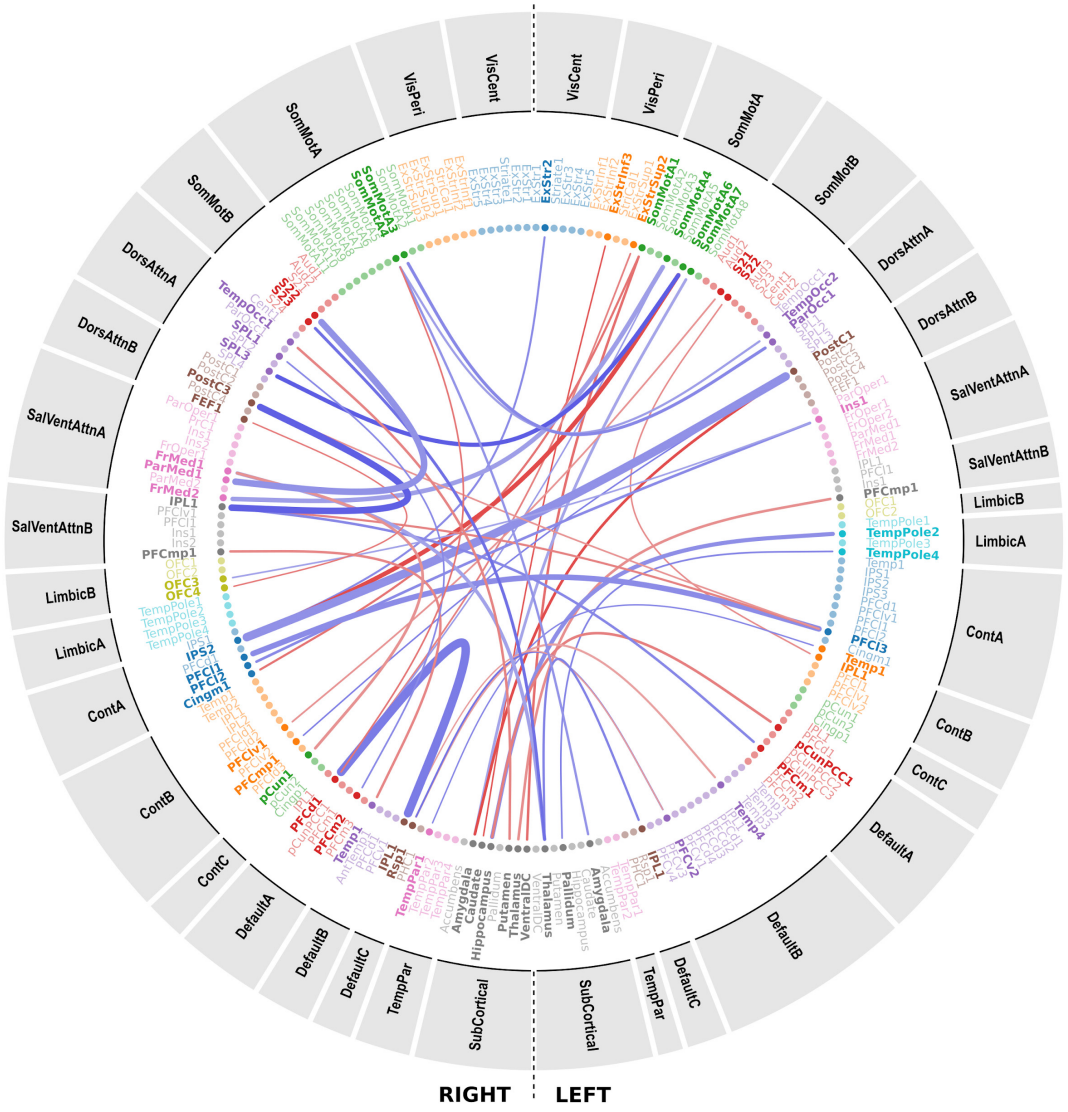


Figure S18. Enlarged edge loadings of $CV_{3|D}$ and $CV_{5|R}$. Edges that load strongly negatively on $u_{3|D}$ and $u_{5|R}$ are depicted in blue. Given baseline differences in the strength of the connectivity patterns, connectivity maps were thresholded at |r| > 0.2 for the discovery sample and at |r| > 0.15 for the replication sample. Next only edges that loaded highly positively or negatively in both datasets were retained for this figure.

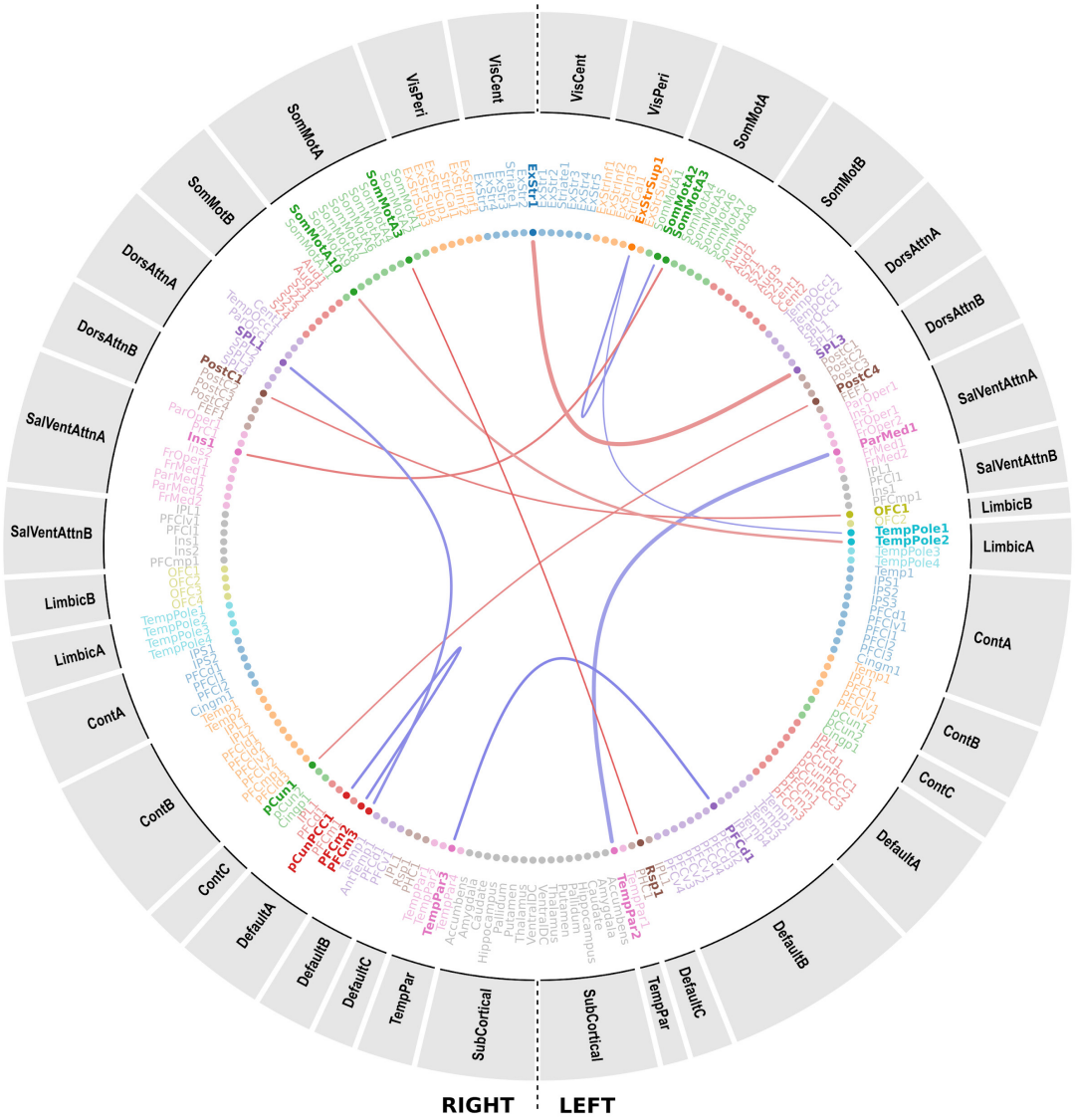


Figure S19. Enlarged edge loadings of $CV_{4|D}$ and $CV_{4|R}$. Depicted edges load strongly positively (red) or strongly negatively (blue) on $u_{4|D}$ and $u_{4|R}$. Given baseline differences in the strength of the connectivity patterns between cohorts, connectivity maps were thresholded at $|r| > 0.2$ for the discovery sample and at $|r| > 0.15$ for the replication sample. Next only edges that loaded highly positively or negatively in both datasets were retained for this figure.

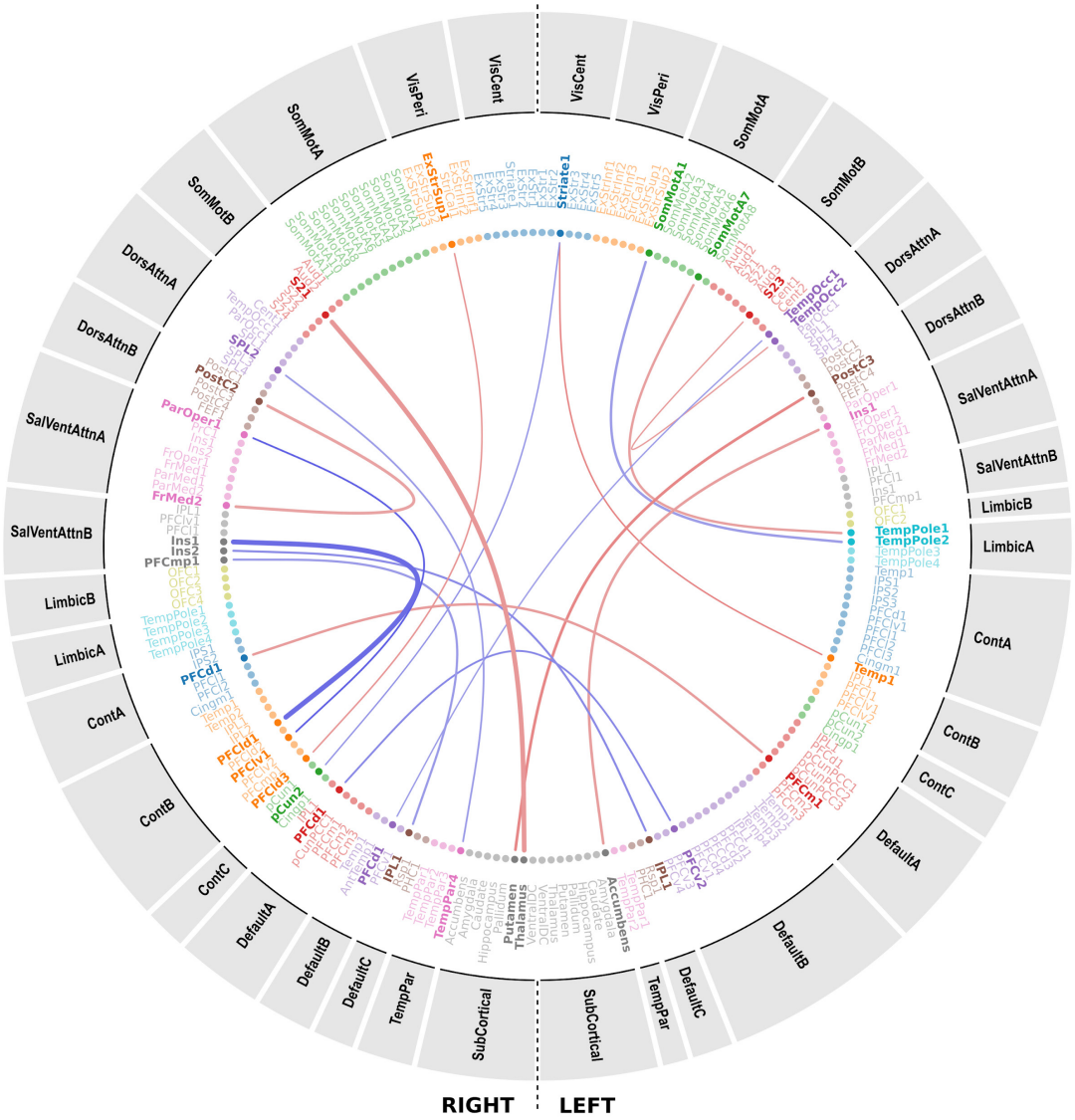


Figure S20. Enlarged edge loadings of $CV_{7|D}$ and $CV_{3|R}$. The circle plot depicts edges in red that load strongly positively (red) or strongly negatively (blue) on $u_{7|D}$ and $u_{3|R}$. Given baseline differences in the strength of the connectivity patterns between cohorts, connectivity maps were thresholded at $|r| > 0.2$ for the discovery sample and at $|r| > 0.15$ for the replication sample. Next only edges that loaded highly positively or negatively in both datasets were retained for this figure.

Supplemental References

1. Alexander LM, Escalera J, Ai L, Andreotti C, Febre K, Mangone A, *et al.* (2017): An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci Data* 4: 170181.
2. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, *et al.* (1997): Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial Reliability and Validity Data. *J Am Acad Child Adolesc Psychiatry* 36: 980–988.
3. Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, Kaufman J, Neer SM (1997): The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale Construction and Psychometric Characteristics. *J Am Acad Child Adolesc Psychiatry* 36: 545–553.
4. Stringaris A, Goodman R, Ferdinando S, Razdan V, Muhrer E, Leibenluft E, Brotman MA (2012): The Affective Reactivity Index: a concise irritability scale for clinical and research settings: The Affective Reactivity Index. *J Child Psychol Psychiatry* 53: 1109–1117.
5. Conners CK, Pitkanen J, Rzepa SR (2011): Conners 3rd Edition (Conners 3; Conners 2008). In: Kreutzer JS, DeLuca J, Caplan B, editors. *Encyclopedia of Clinical Neuropsychology*. New York, NY: Springer New York, pp 675–678.
6. Achenbach TM, Rescorla LA (2001): *Manual for the ASEBA School-Age Forms & Profiles: An Integrated System of Multi-Informant Assessment*. Burlington: University of Vermont: Research Center for Children, Youth and Families.
7. Esteban O, Birman D, Schaer M, Koyejo OO, Poldrack RA, Gorgolewski KJ (2017): MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites ((B. C. Bernhardt, editor)). *PLOS ONE* 12: e0184661.
8. 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.
9. 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
10. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, *et al.* (2004): Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23: S208–S219.
11. Greve DN, Fischl B (2009): Accurate and robust brain image alignment using boundary-

based registration. *NeuroImage* 48: 63–72.

12. Behzadi Y, Restom K, Liau J, Liu TT (2007): A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* 37: 90–101.
13. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE (2014): Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* 84: 320–341.
14. Pruim RHR, Mennes M, van Rooij D, Llera A, Buitelaar JK, Beckmann CF (2015): ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage* 112: 267–277.
15. Ciric R, Wolf DH, Power JD, Roalf DR, Baum GL, Ruparel K, *et al.* (2017): Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *NeuroImage* 154: 174–187.
16. Parkes L, Fulcher B, Yücel M, Fornito A (2018): An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *NeuroImage* 171: 415–436.
17. DuPre E, Gonzalez-Castillo J, Handwerker D, Markello R, Salo T, Whitaker K (2019): Tedana: Robust and extensible software for multi-echo denoising. presented at the OHBM, Rome, Italy.
18. Ségonne F, Dale AM, Busa E, Glessner M, Salat D, Hahn HK, Fischl B (2004): A hybrid approach to the skull stripping problem in MRI. *NeuroImage* 22: 1060–1075.
19. Dale AM, Sereno MI (1993): Improved Localizadon of Cortical Activity by Combining EEG and MEG with MRI Cortical Surface Reconstruction: A Linear Approach. *J Cogn Neurosci* 5: 162–176.
20. Fischl B, Liu A, Dale AM (2001): Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Trans Med Imaging* 20: 70–80.
21. Segonne F, Pacheco J, Fischl B (2007): Geometrically Accurate Topology-Correction of Cortical Surfaces Using Nonseparating Loops. *IEEE Trans Med Imaging* 26: 518–529.
22. Fischl B, Dale AM (2000): Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci* 97: 11050–11055.
23. Fischl B, Sereno MI, Dale AM (1999): Cortical Surface-Based Analysis. *NeuroImage* 9: 195–207.
24. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, *et al.* (2018): Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex* 28: 3095–3114.
25. Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, *et al.* (2011): The organization of the human cerebral cortex estimated by intrinsic functional

connectivity. *J Neurophysiol* 106: 1125–1165.

26. Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TEJ, Glasser MF, *et al.* (2015): A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat Neurosci* 18: 1565–1567.
27. Winkler AM, Renaud O, Smith SM, Nichols TE (2020): Permutation Inference for Canonical Correlation Analysis. *NeuroImage* 117065.
28. Sui J, Adali T, Pearlson G, Yang H, Sponheim SR, White T, Calhoun VD (2010): A CCA+ICA based model for multi-task brain imaging data fusion and its application to schizophrenia. *NeuroImage* 51: 123–134.
29. Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, *et al.* (2016): Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci* 19: 1523–1536.
30. Alnæs D, Kaufmann T, Marquand AF, Smith SM, Westlye LT (2019): *Patterns of Socio-Cognitive Stratification and Perinatal Risk in the Child Brain*. Neuroscience. https://doi.org/10.1101/839969
31. Hyvärinen A (1999): Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Trans Neural Netw* 10: 626–634.
32. Zarzoso V, Comon P (2010): Robust Independent Component Analysis by Iterative Maximization of the Kurtosis Contrast With Algebraic Optimal Step Size. *IEEE Trans Neural Netw* 21: 248–261.
33. Stewart D, Love W (1968): A general canonical correlation index. *Psychol Bull* 70: 160–163.
34. Xia CH, Ma Z, Ciric R, Gu S, Betzel RF, Kaczkurkin AN, et al. (2018): Linked dimensions of psychopathology and connectivity in functional brain networks. *Nat Commun* 9: 3003.