

1 **Multivariate Associations Between Structural Brain Networks, Genetics,**
2 **Environments, and Cognitive-Psychopathological Traits in Children**

3 Jungwoo Seo¹, Eunji Lee², Bo-Gyeom Kim², Gakyung Kim¹, Yoonjung Yoonie

4 Joo^{2,3,4}, Jiook Cha^{1,2,5,6}

- 5 1. Department of Brain and Cognitive Sciences, College of Natural Sciences,
6 Seoul National University, Seoul, 08826, South Korea
- 7 2. Department of Psychology, College of Social Sciences, Seoul National
8 University, Seoul, 08826, South Korea
- 9 3. Department of Digital Health, Samsung Advanced Institute for Health Sciences
10 & Technology (SAIHST), Sungkyunkwan University, Seoul, 06355, South
11 Korea
- 12 4. Research Institute for Future Medicine, Samsung Medical Center, Seoul,
13 06351, South Korea
- 14 5. Institute of Psychological Science, Seoul National University, Seoul, 08826,
15 South Korea
- 16 6. Graduate School of Artificial Intelligence, Seoul National University, Seoul,
17 08826, South Korea

18

19

20 **Correspondence to:** Jiook Cha, PhD

21 Gwanak-ro 1, Building 16, Suite M512, Gwanakgu, Seoul, 08826, South
22 Korea, connectome@snu.ac.kr

23 **Short title:** Children's brain network, gene, environment, and cognitive-behavior

24 **Keywords:** structural brain network, genetic factor, environmental factor, cognitive-
25 behavior, preadolescent, multivariate analysis

26

27 **Abstract**

28 Brain development in childhood is shaped by complex interactions between genetic
29 predispositions, environmental influences, and neural connectivity, yet how these
30 factors jointly contribute to cognitive and mental health outcomes remains unclear.
31 Structural brain networks, quantified through graph-theoretic measures, have been
32 linked to cognition and psychiatric risk, but the extent to which genetic architecture
33 and environmental exposures shape these networks, and whether brain networks
34 mediate these influences, is not well understood. Here we show that genetic
35 predispositions related to cognitive ability and socioeconomic status (SES) exhibit
36 the strongest covariation with structural brain network topology in children. Using
37 sparse canonical correlation analysis (SCCA) on ABCD Study data ($N = 10,343$), we
38 identified robust associations between brain network properties, polygenic scores for
39 cognition, SES indicators, and cognitive-psychopathological phenotypes. Mediation
40 analysis further demonstrated that structural brain networks partially mediate the
41 influence of genetic and environmental factors on cognitive performance and mental
42 health outcomes, suggesting that neurodevelopmental trajectories may be shaped
43 by both genetic liability and modifiable environmental conditions. These findings
44 provide empirical support for a multivariate, systems-level perspective on brain
45 development and cognitive-psychopathological variation in youth. By elucidating
46 shared neural substrates underlying genetic and environmental influences, this work
47 advances our understanding of brain network development and highlights potential
48 pathways for individualized interventions and predictive modeling in developmental
49 psychiatry and neuroscience.

50 Introduction

51 Childhood and adolescence are critical periods for brain development
52 (Bethlehem et al., 2022). Proper brain development during this time is vital for
53 cognitive and behavioral maturation (Bunge & Wright, 2007; Luna et al., 2010) and
54 mental health (Fornito et al., 2015; Paus et al., 2008). This development is shaped by
55 a complex interplay between genetic predispositions, environmental influences, and
56 brain network dynamics, yet the precise mechanisms underlying these relationships
57 remain poorly understood. Therefore, understanding the connections between the
58 brain, cognitive-behavioral traits in children, and the impact of genetics and the
59 environment on brain development is crucial in developmental and clinical
60 neuroscience.

61 The Adolescent Brain and Cognitive Development (ABCD) study provides an
62 unprecedented opportunity to examine these interrelations in a large, representative
63 sample (Jernigan et al., 2018). Leveraging multimodal neuroimaging, genome-wide
64 polygenic scores (PGSs), and extensive environmental and behavioral assessments,
65 the ABCD dataset enables a more comprehensive modeling of developmental
66 trajectories than previously possible.

67 The brain is best understood as a complex, interconnected network, rather
68 than a collection of isolated regions. Graph-theoretic approaches to brain connectivity
69 provide powerful tools to quantify structural network properties such as integration
70 (e.g., global efficiency), segregation (e.g., modularity), and centrality (e.g., degree
71 centrality) (Rubinov & Sporns, 2010). Diffusion MRI and tractography allow for the
72 reconstruction of white matter networks, enabling researchers to examine how these

73 structural connections support cognition (Jeurissen et al., 2019; Sotiropoulos &
74 Zalesky, 2019). Notably, individual differences in brain network organization have been
75 linked to cognitive abilities (Bathelt, Gathercole, Butterfield, et al., 2018; Kim et al.,
76 2016; Koenis et al., 2015; Ma et al., 2017; Suprano et al., 2020), psychiatric risk
77 (Alexander-Bloch et al., 2010; Collin et al., 2017; Rudie et al., 2012), and
78 environmental influences such as socioeconomic status (Kim et al., 2019; Tooley et
79 al., 2020). However, how genetic and environmental factors jointly shape structural
80 brain networks, and whether these networks mediate their influence on cognitive and
81 psychiatric outcomes, remains poorly understood.

82 Twin and heritability studies suggest that brain network properties are, at least
83 in part, genetically influenced, with estimates of heritability for measures like global
84 efficiency ranging from 25% to 70% (Koenis et al., 2015; van den Heuvel et al., 2013).
85 Yet, the specific genetic variants shaping these structural networks—particularly in the
86 context of cognitive and psychiatric outcomes—remain poorly characterized.
87 Investigating this gene-brain-behavior interplay is essential for understanding how
88 biological predispositions manifest in cognitive and mental health outcomes during
89 development.

90 Despite valuable insights from previous studies, most prior research has
91 employed univariate approaches, limiting the ability to capture the multivariate
92 complexity of the gene/environment-brain-behavior relationship. To address this
93 limitation, canonical correlation analysis (CCA) is useful for investigating holistic
94 relationships underlying a set of variables. CCA models multivariate associations
95 linking sets of variables from two or more domains by maximizing the canonical
96 correlation between them (Hotelling, 1936; Wang et al., 2020). Indeed, CCA has been

97 widely applied in studying links between brain connectivity, cognitive function, genetics,
98 and environmental factors (Alnaes et al., 2020; Fernandez-Cabello et al., 2022;
99 Modabbernia et al., 2021; Smith et al., 2015; Wang et al., 2020). However, traditional
100 CCA methods may struggle with high-dimensional datasets, necessitating more
101 advanced approaches. To further address these limitations, we apply sparse canonical
102 correlation analysis (SCCA) (Witten et al., 2009), a multivariate method that identifies
103 maximally covarying patterns between genetic factors, brain network properties, and
104 behavioral traits while reducing dimensionality and enhancing interpretability.

105 Utilizing the ABCD dataset's expansive genetic-environmental-neuroimaging-
106 behavioral data, this study aims to answer two fundamental questions. Firstly, "How
107 do genetic predispositions/environmental influences shape structural brain network
108 properties in children?"; Secondly, "Do brain network properties mediate the influence
109 of genetic and environmental factors on cognitive or psychiatric phenotypes?"

110 By integrating genetics, environmental, neuroimaging, and behavioral
111 assessments in a large, developmentally significant cohort, this study seeks to
112 elucidate the complex gene-environmental-brain-behavior pathways that shape
113 cognitive and mental health outcomes. This research contributes to a more
114 comprehensive understanding of the biological and environmental foundations of brain
115 development and lays the groundwork for future interventions aimed at optimizing
116 cognitive and psychological well-being in youth.

117 **Methods and Materials**

118 **ABCD participants**

119 We used genetic, neuroimaging, environmental, and phenotypic data from
120 the Adolescent Brain Cognitive Development (ABCD) study (<http://abcdstudy.org>),
121 specifically from release 2.0 for genetic and neuroimaging data, release 3.0 for
122 genetic ancestry information, and release 5.1 for environmental and phenotypic data.
123 The ABCD study, the largest longitudinal investigation of brain development and
124 child health in the United States, recruited multiethnic children (N=11,875) aged 9-10
125 years from 21 research sites. The self-reported ethnicities of participants included
126 52.3% White, 20.3% Hispanic, 14.7% Black, and 12.5% Asian and others. All
127 participants provided informed assent, and their parents or legal guardians provided
128 informed consent before participating in the study.

129

130 **Genotype Data and Polygenic Scores**

131 The genotype data used in this study were obtained from the ABCD cohort,
132 with DNA samples genotyped at Rutgers University Cell and DNA Repository
133 (RUCDR) using the Affymetrix NIDA Smoke Screen Array. Standard quality control
134 (QC) procedures were applied to remove variants with low genotype call rate and
135 minor allele frequency (MAF). Genotype imputation was performed using the
136 Michigan Imputation Server (Das et al., 2016) with the 1000 Genomes reference
137 panel (Genomes Project et al., 2015). We further filtered out imputed variants that
138 did not meet our QC criteria. To account for population structure and relatedness,
139 principal component analysis (PCA) and kinship-based filtering were used to exclude
140 close relatives (Conomos et al., 2015; Conomos et al., 2016).

141 Polygenic scores (PGSs) for 30 cognitive, psychiatric, and behavioral traits
142 were calculated based on genome-wide association study (GWAS) summary
143 statistics. These included PGSs for cognitive performance, educational attainment,
144 schizophrenia, bipolar disorder, depression, insomnia, body mass index, and
145 automobile speed propensity, among others. The PGSs were derived using PRS-CS
146 (Ge et al., 2019), a Bayesian regression approach, with optimal hyperparameter

147 selection.

148 The genotype data and polygenic scores used in this study are identical to
149 those reported in (Joo et al., 2024). A complete list of the 30 polygenic scores, along
150 with further methodological details on genotype data processing, GWAS sources,
151 ancestry-based adjustments, and validation procedures can be found in the
152 **Supplementary Materials** or (Joo et al., 2024).

153

154 **Environmental Factors**

155 To investigate the relationship between children's brain network properties
156 and their environment, we analyzed 56 variables representing various environmental
157 aspects. These variables included indicators of family and neighborhood
158 socioeconomic status, such as family income, parental education level, marital
159 status, and neighborhood deprivation based on the Area Deprivation Index (ADI).
160 Additionally, variables related to prenatal substance exposure (e.g., tobacco, alcohol,
161 cocaine, and marijuana), as well as parental and family factors such as parental
162 acceptance and family conflict, were included.

163 To maximize statistical power, we imputed missing data to include as many
164 samples as possible in the analysis. To account for uncertainty in imputing missing
165 values and to improve the accuracy and reliability of the imputed values, we
166 employed the Multiple Imputation by Chained Equations (MICE) method (Van
167 Buuren & Groothuis-Oudshoorn, 2011). We set the number of imputation iterations to
168 40 to ensure stable convergence. All imputation procedures were conducted using
169 the `statsmodels` package in Python. To ensure that our analysis was robust to

170 imputation, we also conducted analyses using only the complete data without any
171 missing values and included these results as supplementary material.

172

173 **Phenotype Data**

174 To explore the relationship between children's brain network properties and
175 their cognitive ability, mental well-being, and physical health, we examined 86
176 phenotype variables. To evaluate cognitive ability, we examined the NIH Toolbox
177 measurements, which included fluid, crystallized, and overall cognition scores, along
178 with domain-specific task scores for episodic memory, executive function, language,
179 and processing speed (Weintraub et al., 2013). For assessing the children's mental
180 health, our analysis encompassed a broad range of emotional and behavioral
181 measurements, including instruments such as the Child Behavioral Checklist
182 (CBCL), the Kiddie-Structured Assessment for Affective Disorders and Schizophrenia
183 for DSM-5 (KSADS-5), and the Prodromal Psychosis Scale (PPS). We also included
184 measures of behavioral tendencies, such as behavioral inhibition and activation
185 (BIS/BAS), which assess avoidance behaviors, reward sensitivity, and behavioral
186 control; impulsivity traits (UPPS), which measure dimensions such as urgency,
187 perseverance, and sensation seeking; and sleep-related problems, which provide
188 insights into sleep patterns, quality, and related disorders. For the children's physical
189 health, we included data such as physical activity. To maximize statistical power,
190 missing values in the phenotype data were imputed using the Multiple Imputation by
191 Chained Equations method (Van Buuren & Groothuis-Oudshoorn, 2011).

192

193 **Structural Brain Network Construction**

194 Detailed procedures for acquiring and preprocessing MRI data are described
195 in (Kim et al., 2022). In brief, we used structural and diffusion MRI data acquired by
196 the ABCD study (Casey et al., 2018) from data release 2.0. The preprocessing steps,
197 as detailed in (Kim et al., 2022), included eddy current and head motion correction,
198 diffusion gradient adjustment, and various distortion corrections. Quality control (QC)
199 was performed using freesurfer QC metric (fsqc_qc) and raw dMRI QC metric
200 (iqc_dmri_1_qc_score). To estimate brain structural networks from neuroimaging,
201 individual connectome data was generated. This was achieved by applying MRtrix3
202 (Tournier et al., 2019) to the preprocessed dMRI data to estimate whole-brain white
203 matter tracts and generate individualized connectomes. Probabilistic tractography
204 was performed using constrained-spherical deconvolution (CSD) (Calamante et al.,
205 2010; Tournier et al., 2007) with random seeding across the brain and target
206 streamline counts of 20 million. Initial tractograms were filtered using spherical-
207 deconvolution informed filtering (2:1 ratio)(Smith et al., 2013), resulting in a final
208 streamline count of 10 million. An 84x84 whole-brain connectome matrix was
209 generated for each participant using the T1-based parcellation and segmentation
210 from FreeSurfer with Desikan-Killiany atlas (Desikan et al., 2006)(68 nodes for the
211 cortical regions and 16 nodes for the subcortical regions). This approach ensured
212 that individual participants' connectomes were based on their neuroanatomy. The
213 computation was conducted on supercomputers at Argonne Leadership Computing
214 Facility Theta and Texas Advanced Computing Center Stampede2.

215

216 **Brain Network Measures (BNMs)**

217 We used the connectome matrix to construct an undirect weighted graph
218 representing the structural brain network. Nodes and edges in the graph represent
219 parcellated gray matter regions and connections between them, respectively.
220 Connection strength was quantified by the streamline counts. To account for the
221 potential false positive connections generated by probabilistic tractography and their
222 impact on network topology, we eliminated extremely weak connections (streamline
223 counts less than 3). After thresholding, we excluded individuals with at least one
224 isolated node, assuming all brain regions are communicable via at least one path.

225 We calculated 13 different types of brain network measurements (BNMs)
226 representing different aspects of brain network's property (Rubinov & Sporns, 2010;
227 van den Heuvel & Sporns, 2011). We calculated eight global graph metrics (including
228 network density, modularity, normalized modularity, normalized average clustering
229 coefficient, normalized characteristic path length, global efficiency, normalized global
230 efficiency, small worldness) and five nodal graph metrics (including degree, strength,
231 clustering coefficient, betweenness centrality, nodal efficiency) to represent brain
232 network's global and regional properties. All graph measures were calculated using
233 the package Brain Connectivity Toolbox (<https://sites.google.com/site/bctnet/>).

234

235 **Statistical analysis**

236 *Sparse Canonical Correlation Analysis*

237 To examine a latent mode of covariation between structural brain network
238 properties and various polygenic scores, environmental factors, and phenotypic
239 outcomes, we used sparse canonical correlation analysis (Witten et al., 2009)
240 between brain network measures and three types of non-imaging data (i.e., PGSs,
241 environmental variables, phenotype variables) separately. While traditional CCA can
242 be effective, it often suffers from overfitting and interpretability issues in high-
243 dimensional datasets. SCCA, with its L1 regularization, addresses these issues by
244 producing sparse solutions that enhance interpretability and reduce overfitting.
245 Although our dataset is not high-dimensional enough to make traditional CCA
246 infeasible, we chose SCCA to identify interpretable patterns and ensure robust
247 results.

248 The most popular algorithm for sparse canonical correlation analysis is
249 penalized matrix decomposition (PMD)(Witten et al., 2009), which solves
250 optimization problem of below equation for given two sets of data matrix $X_{n \times p}$, $Y_{n \times q}$.
251 (n: sample size; p, q: the number of variables of domain X and Y respectively; u, v:
252 canonical weights of domain X and Y respectively; c1, c2: regularization parameter)

$$253 \quad \max cov(Xu, Yv)$$

$$254 \quad s.t. \quad \|u\|_2 = \|v\|_2 = 1, \quad \|u\|_1 \leq c_1, \quad \|v\|_1 \leq c_2$$

255 To interpret Witten's sparse canonical correlation analysis as correlation
256 maximization, we need to assume covariance matrices $X^T X$, $Y^T Y$ are identity
257 matrices (Witten et al., 2009). But in our study with the high dimensional brain
258 datasets, the assumption is hardly satisfied. For this reason, we interpreted Witten's

259 sparse canonical correlation analysis as a maximizing covariance algorithm between
260 two sets of variables rather than maximizing correlation.

261 To test the generalizability of the sparse canonical correlation analysis
262 results, we split the dataset into a training and test set. To reduce ABCD site-
263 sensitive bias, we performed a stratified train (80%) - test (20%) split based on site
264 ID. For the sparse canonical correlation analysis with PGS and brain network
265 measures, we only used participants classified as genetically European ancestry to
266 control genetic confounding effects as the main analysis. To ensure that these
267 analyses generalize to a multiethnic dataset, we conducted additional analyses on
268 the multiethnic dataset. In the multiethnic analysis, genetic ancestry was included as
269 an additional covariate. **Table 1 and Supplementary Table 1** summarize the
270 demographic information of the samples included in main and supplementary
271 analysis, respectively.

272 We attempted to control for potential confounding effects from age, sex, self-
273 reported race/ethnicity, ABCD study site, and handedness. Similar to (Modabbernia
274 et al., 2021), we controlled these potential confounding effects by regressing out the
275 variance explained by age, sex, age * sex, age², age² * sex, self-reported
276 race/ethnicity, ABCD study site, and handedness from all variables prior to
277 performing SCCA analysis. For binary variables such as KSADS diagnosis, we used
278 logistic regression to regress out these effects. The residualized data was then used
279 as input for sparse canonical correlation analysis (SCCA). We used `SCCA_PMD`
280 function from the Python 'cca-zoo' package. To ensure that the main findings are
281 robust and not driven by the selection of specific covariates, we performed additional

282 analyses using a reduced set of covariates (age, sex, self-reported race/ethnicity,
283 ABCD study site, and handedness).

284 We selected optimal L1 regularization parameters from 5-fold cross validation
285 searching from 0.1 to 1 with a step size of 0.05 for both X and Y variables
286 respectively. The optimal L1 parameter combination was selected to maximize the
287 covariance of validation set between canonical variates of the first component.

288 For each sparse canonical correlation analysis, we extracted five modes of
289 covariance. To examine the statistical significance of each mode, we used a
290 permutation test. By randomly shuffling the rows of one dataset and remaining the
291 other, we generated 5,000 permutation sets. The p-value of each component was
292 calculated based on the number of permutation sets having greater covariance than
293 that obtained from the original dataset, and FDR-correction was done within each
294 CCA.

295
$$p_{uncorrected} = \frac{N_{null\ cov > cov}}{N_{null}}$$

296 Selected variables and their loading depend on the input sample. To find
297 variables reliably related to each mode, we used bootstrap resampling. We randomly
298 resampled 5,000 times with replacement and assessed the 95% confidence interval
299 of each variable's loading and how consistently it was selected. We interpreted the
300 significant modes based on loading patterns of variables whose 95% confidence
301 interval of loading does not cross zero (Xia et al., 2018) and selected more
302 frequently than expected by chance (i.e., more frequently selected than expected by
303 binomial distribution). Because sparse canonical correlation analysis with bootstrap
304 sample may change the order of components (axis rotation) and signs (reflection)

305 (Misic et al., 2016; Xia et al., 2018), the re-alignment procedure is needed to
306 estimate confidence interval of loading properly. We matched the components and
307 signs based on cosine similarity of weight vectors obtained from original dataset and
308 bootstrap sample. To assess the reproducibility of the findings, we applied the model
309 to the held-out test set and estimated significance of each mode through the
310 permutation test.

311 *Mediation Analysis*

312 After performing sparse canonical correlation analysis (SCCA), we tested
313 whether brain network properties mediate the influence of genetic and environmental
314 factors on cognitive and psychiatric phenotypes. Since SCCA identifies latent modes
315 of covariation between brain network properties and multiple domains, the SCCA-
316 derived variates serve as optimal summary representations that capture the
317 dominant axes of covariation among brain network properties, genetic
318 predispositions, environmental influences, and phenotypic traits. We used these
319 summary scores in mediation analysis to test the hypothesized gene/environment–
320 brain network–phenotype pathway while reducing high-dimensional data into
321 interpretable components.

322 We used the same covariates as in the SCCA, which include age, sex, age *
323 sex, age², age² * sex, self-reported race/ethnicity, ABCD study site, and
324 handedness. The two-sided p-values for each path were estimated from 500
325 bootstrap samples using the `mediation_analysis` function from the Python 'pingouin'
326 package.

327 Results

328 Using sparse canonical correlation analysis (SCCA), we investigated
329 relationships between structural brain network properties and three domains:
330 genome-wide polygenic scores (PGSs), environmental factors, and phenotypic
331 outcomes. **Figure 1** presents only variables whose 95% confidence intervals for
332 loadings do not cross zero, ensuring reliable positive or negative associations.

333 Covariation between polygenic scores and structural brain network properties

334 Our analysis identified a statistically significant mode of covariation between 30
335 genome-wide polygenic scores and brain network measures (PGS-BNM mode 1, $p =$
336 0.001 , $cov = 0.895$, $r = 0.143$, FDR corrected).

337 In this mode of PGSs (**Figure 1A**), polygenic scores reflecting cognitive ability
338 (e.g., IQ and educational attainment) showed the strongest positive loadings, while the
339 polygenic score for BMI showed a negative loading. Additionally, polygenic scores for
340 automobile speed propensity and cannabis use had moderate positive loadings,
341 whereas those for insomnia showed moderate negative loadings. Given these
342 dominant loading patterns, this mode represents variance along the ‘Cognitive-
343 Obesity Genetic Axis,’ primarily reflecting genetic traits related to cognitive ability.
344 Loadings for polygenic scores not shown in (**Figure 1A**) are provided in
345 (**Supplementary Figure 1**).

346 Higher values in this mode were linked to increased nodal efficiency in the
347 temporal gyri, betweenness centrality in the supramarginal and post-central gyri, and
348 higher clustering coefficients in the temporal, parietal, central, and inferior frontal gyri
349 (**Figure 1B**). Conversely, higher values in this mode were associated with lower

350 connectivity in the cingulate cortex, insula, and subcortical regions, along with reduced
351 nodal efficiency in the insula and subcortical regions. Regarding global brain network
352 measures, density (-0.613) and global efficiency (-0.421) showed negative loadings,
353 while the normalized average clustering coefficient (0.542) showed a positive loading.

354 Results from the multi-ethnic analysis were consistent with those observed in
355 the European-only analysis (**Supplementary Figure 2**). When we controlled for
356 confounding effects using a simplified set of covariates, the results remained largely
357 unchanged. However, when the SCCA model trained on the training set was applied
358 to the test set, the p-value was marginally significant ($p = 0.0534$).

359

360 **Covariation between environmental factors and structural brain network** 361 **properties**

362 Among the five modes analyzed, only the first mode was statistically
363 significant and generalized to the hold-out test set (ENV-BNM mode 1, $p < 0.001$,
364 $cov = 1.251$, $cor = 0.146$, FDR corrected). The environmental component of this
365 mode predominantly captured variance related to socioeconomic status (SES)
366 (**Figure 1C**).

367 Notably, neighborhood-level SES variables, such as ADI median family
368 income and ADI education level, showed the strongest positive loadings. Similarly,
369 family-level SES variables, including household income, also showed positive
370 associations, though their loadings were slightly lower than those of neighborhood-
371 level SES measures. Variables reflecting neighborhood-level socioeconomic

372 deprivation, such as ADI poverty indices, showed strong negative loadings.
373 Additionally, prenatal substance exposures—such as tobacco exposure during
374 pregnancy—showed moderate negative associations with this component. These
375 results suggest that this component reflects a spectrum of socioeconomic status
376 from socioeconomic advantage to disadvantage.

377 On the brain network side (**Figure 1D**), connection strength and nodal
378 efficiency in the temporal gyrus showed positive loadings, and connection strength,
379 betweenness centrality, and clustering coefficient in regions such as the cingulate
380 cortex and precuneus showed negative loadings. Regarding global brain network
381 measures, normalized global efficiency (0.360) showed positive loading.

382 These results were robust across different sets of covariates. When missing
383 values were handled by dropping incomplete cases rather than through imputation,
384 the environmental component remained consistent. The general loading patterns of
385 brain network properties were also similar; but, the overall magnitude of the loadings
386 was slightly lower than in the main analysis, possibly due to reduced statistical power
387 (**Supplementary Figure 4**).

388

389 **Covariation between phenotypes and structural brain network properties**

390 Among the five modes of covariation identified by SCCA, the first two modes
391 were statistically significant and generalized to the hold-out test sets (Pheno-BNM
392 mode 1: $p < 0.001$, $cov = 1.421$, $r = 0.143$; Pheno-BNM mode 2: $p < 0.001$, $cov = 0.923$,
393 $r = 0.123$, all p-values were FDR corrected).

394 The first mode captured covariation between brain network properties and an
395 integrated measure of cognitive ability and psychopathology in children (**Figure 1E**).
396 Cognitive ability-related scores (e.g., NIH Toolbox scores) showed positive loadings,
397 while psychopathology-related scores (e.g., CBCL scores) showed negative
398 loadings. These loading patterns suggest that the first phenotype mode reflects
399 variation along the cognitive ability-psychopathology axis.

400 Brain network properties associated with this mode showed positive loadings
401 for nodal efficiency and betweenness centrality in the temporal and parietal cortices.
402 In contrast, connection strength in the insula, cingulate, precuneus, and subcortical
403 regions showed negative loadings. Additionally, nodal efficiency in the insula and
404 subcortical regions was negatively associated with this mode (**Figure 1F**). No
405 significant global brain network measures were identified for this mode.

406 In phenotype mode 2 (**Figure 2**), positive loadings were observed for
407 standing height, being born prematurely, having hearing or vision issues, and
408 experiencing obstetric complications. Although the uncertainty in loading estimation
409 is considerable, abnormal behavior (CBCL scores) also exhibited the highest positive
410 loading in this mode (**Supplementary Figure 7**). On the brain network properties
411 side, lower connection strength and nodal efficiency were observed in the precentral
412 and postcentral gyri, superior frontal gyrus, and thalamus. Degree in the middle
413 temporal gyrus, superior frontal gyrus, and postcentral gyrus also tended to be lower.
414 Among global brain network metrics, normalized modularity (-0.195), global
415 efficiency (-0.343), and normalized global efficiency (-0.487) showed negative
416 loadings, while modularity (0.534) showed positive loading.

417 The findings remained stable across different covariate sets. When missing
418 values were handled by excluding incomplete cases instead of using imputation, the
419 mode 1 results remained consistent (**Supplementary Figure 6**). However, in mode
420 2, unlike the imputed case, only height, weight, and BMI exhibited reliable positive
421 loadings, while CBCL, despite its high estimation uncertainty, showed a strong
422 negative loading (**Supplementary Figure 8**).

423

424 **Shared Covariation Patterns in Brain Network Properties**

425 Our SCCA identified certain brain network properties that showed similar
426 covariation patterns across genetic factors related to cognitive ability, socioeconomic
427 status, and phenotypes of cognitive ability-psychopathology. (**Figure 1G**) presents
428 the brain network properties that shared common loading patterns with the primary
429 SCCA modes for polygenic scores (PGS mode 1), environmental factors (ENV mode
430 1), and phenotypes (Pheno mode 1). Across these domains, brain network measures
431 such as nodal efficiency in the temporal gyrus consistently showed positive loadings,
432 while measures like connection strength in the posterior cingulate and subcortical
433 regions consistently showed negative loadings.

434

435 **Mediation Analysis**

436 The observed covariation patterns in brain network properties across genetic,
437 environmental, and phenotypic domains led us to hypothesize that structural brain
438 networks mediate two key relationships: (1) between genetic factors related to

439 cognitive ability and the cognitive ability-psychopathology phenotype, and (2)
440 between socioeconomic status and the cognitive ability-psychopathology phenotype.
441 To test these hypotheses, we conducted mediation analyses using SCCA-derived
442 summary scores from each domain.

443 Our analysis revealed significant mediation effects in both cases. Structural
444 brain network properties partially mediated the relationship between polygenic
445 scores for cognitive ability and the cognitive ability-psychopathology phenotype
446 (indirect effect = 0.023, $p < 0.001$; **Figure 3A**), as well as between socioeconomic
447 status and the cognitive ability-psychopathology phenotype (indirect effect = 0.015, p
448 < 0.001 ; **Figure 3B**).

449

450 **Discussion**

451 This study investigated how structural brain network properties covary with
452 genetic, environmental, and phenotypic factors in 9–10-year-old children and
453 whether these properties mediate genetic and environmental effects on cognitive-
454 behavioral outcomes. By integrating polygenic scores (PGS), environmental
455 variables, and cognitive-behavioral traits using sparse canonical correlation analysis
456 (SCCA), we identified shared covariation patterns in brain network properties across
457 these domains. Previous studies have demonstrated that structural brain network
458 properties are heritable (Koenis et al., 2015; van den Heuvel et al., 2013), but the
459 specific genetic contributions to brain network organization in children remain
460 underexplored. Our findings address this gap by providing evidence that cognitive
461 ability-related genetic factors, socioeconomic status, and cognitive-

462 psychopathological phenotypes are key determinants of structural brain network
463 variations in preadolescents. Moreover, mediation analyses reveal that structural
464 brain network properties serve as intermediaries between genetic/environmental
465 influences and cognitive-psychopathological outcomes. This suggests that variations
466 in brain network organization may provide a mechanistic pathway through which
467 early-life genetic and environmental factors contribute to individual differences in
468 cognitive and mental health outcomes.

469 Our results reveal that structural brain network properties exhibit distinct yet
470 overlapping covariation patterns with genetic, environmental, and phenotypic factors.
471 Specifically, brain network measures such as nodal efficiency in the temporal and
472 parietal cortices were consistently associated with cognitive ability-related genetic
473 factors, higher socioeconomic status, and better cognitive performance. Conversely,
474 weaker connectivity in the posterior cingulate, insula, and subcortical regions was
475 commonly linked to genetic risk for lower cognitive ability, socioeconomic
476 disadvantage, and increased psychopathology. Consistent with prior studies (Alnaes
477 et al., 2020; Fernandez-Cabello et al., 2022; Modabbernia et al., 2021; Smith et al.,
478 2015), these findings suggest that children's brain network properties also covary
479 along spectrums of 'positive-negative' factors across genetic, environmental, and
480 phenotypic dimensions. Unlike previous studies that explored covariation between
481 brain features and only some of these domains, our study examined all three factors
482 simultaneously, identifying a common set of brain network properties that covary
483 across genetic, environmental, and phenotypic factors. These findings highlight the
484 potential of these brain networks as key substrates underlying cognitive and mental

485 health disparities, shaped by both genetic predisposition and environmental
486 influences during childhood.

487 Our findings align with established neurodevelopmental models, which
488 suggest that childhood and adolescence are characterized by a shift from
489 subcortical-driven to cortico-cortical-dominated network organization (Baker et al.,
490 2015; Langen et al., 2018; Menon, 2013; Sato et al., 2015). Specifically, favorable
491 genetic and phenotypic traits related to cognitive ability and SES were associated
492 with greater nodal efficiency and betweenness centrality in the temporal and parietal
493 cortices and reduced connectivity in the posterior cingulate, insula, and subcortical
494 regions. These patterns are consistent with established neurodevelopmental
495 trajectories, which involve a progressive strengthening of long-range cortico-cortical
496 connectivity (Hwang et al., 2013; Oldham & Fornito, 2019) and reduction in
497 subcortical connectivity as higher-order networks become more specialized (Baker et
498 al., 2015; Langen et al., 2018; Sato et al., 2015). Interestingly, the associations
499 between genetic, environmental, and phenotypic factors and brain network
500 organization mirror expected patterns of cortical-subcortical reorganization, wherein
501 cortico-cortical integration strengthens while subcortical connectivity decreases.
502 These findings suggest that genetic and environmental influences may shape the
503 timing or pace of this cortical-subcortical transition, potentially accelerating or
504 delaying neurodevelopmental trajectories and influencing individual differences in
505 cognitive and mental health outcomes (Heller et al., 2016).

506 Children with higher cognitive performance tend to have greater nodal
507 efficiency and betweenness centrality in the temporal, parietal, and superior frontal
508 regions—areas crucial for higher-order cognitive functions such as language,

509 semantic processing, abstract reasoning, and working memory (Binder et al., 2009;
510 Culham & Kanwisher, 2001; du Boisgueheneuc et al., 2006; Price, 2012; Visser et
511 al., 2012). This suggests that enhanced network integrity in these regions supports
512 cognitive ability by efficient communication across the brain. Consistent with this, our
513 results support the parieto-frontal integration theory (P-FIT) of intelligence (Basten et
514 al., 2015; Jung & Haier, 2007), which posits that effective information integration
515 across distributed networks, particularly involving the frontal, parietal, and temporal
516 regions, underlies intelligence. Notably, while previous studies have linked global
517 efficiency with cognitive performance (Bathelt, Gathercole, Butterfield, et al., 2018;
518 Kim et al., 2016; Ma et al., 2017), our findings—derived from a largest to date
519 sample—suggest that regional network efficiency, rather than global efficiency, is
520 more strongly associated with cognitive ability. This highlights the importance of
521 region-specific network topologies in understanding neurodevelopmental differences.

522 Although neighborhood-level socioeconomic status (SES) and family-level
523 SES are both associated with children’s brain development, neighborhood-level SES
524 has a unique relationship with brain structure and functional networks, distinct from
525 that of family-level SES (Rakesh et al., 2022; Tooley et al., 2020). In our results, the
526 link between neighborhood-level SES and structural brain network properties was
527 more pronounced than associations with family-level SES.

528 Prenatal substance exposure, particularly to tobacco and marijuana,
529 exhibited moderate negative loadings within the SES-related brain network mode,
530 reinforcing prior findings that lower socioeconomic status is associated with
531 increased prenatal exposure to neurotoxic substances (Gu et al., 2024; Metz et al.,
532 2018; Mravcik et al., 2020). This association likely reflects a complex interplay

533 between socioeconomic adversity and prenatal environmental stressors, both of
534 which have been implicated in shaping neurodevelopmental trajectories (El Marroun
535 et al., 2016; Ross et al., 2015; Thompson et al., 2009). Given that SES-related
536 disparities in brain connectivity may stem from a combination of prenatal exposures,
537 postnatal environments, and genetic predispositions, the observed negative loading
538 may capture broader socioeconomic influences rather than a direct teratogenic effect
539 of prenatal substance exposure. However, disentangling these effects is challenging,
540 as the precise mechanisms through which prenatal exposures contribute to structural
541 brain network alterations remain unclear. Future studies leveraging genome-
542 environment interaction analyses, longitudinal neuroimaging, and causal inference
543 approaches (e.g., Mendelian randomization) are essential to elucidate how prenatal
544 risk factors, SES, and genetic predispositions collectively shape neurodevelopmental
545 outcomes.

546 Phenotype mode 2 presents a complex pattern, linking standing height,
547 perinatal risk factors, and sensorimotor brain networks. The lower network efficiency
548 in the precentral and postcentral gyri, superior frontal gyrus, and thalamus suggests
549 a potential connection to early neurodevelopmental processes, particularly those
550 involved in motor and sensory integration. Given the variability in loading estimates,
551 further research is needed to determine whether these associations reflect specific
552 neurodevelopmental mechanisms or statistical artifacts.

553 Several limitations should be acknowledged. First, the cross-sectional design
554 precludes causal inferences, necessitating longitudinal research to validate the
555 mediating role of brain networks in shaping developmental trajectories of cognitive
556 and mental health. Second, gene-by-environment (G×E) interactions were not

557 explicitly modeled, limiting our ability to assess whether socioeconomic status (SES)
558 moderates genetic influences on brain connectivity. Future research should integrate
559 G×E interaction analyses to better understand how genetic predispositions interact
560 with environmental contexts in shaping neurodevelopment. Finally, although sparse
561 canonical correlation analysis (SCCA) provided a powerful multivariate approach, it
562 assumes linear relationships between brain networks and genetic/environmental
563 factors, which may oversimplify complex neurodevelopmental processes. Future
564 studies should consider non-linear modeling approaches and more sophisticated
565 causal inference methods (e.g., Mendelian randomization, structural equation
566 modeling) to capture the intricate interplay of genes, environments, and brain
567 development.

568 Addressing these gaps will enhance the robustness of future research,
569 ultimately contributing to a more comprehensive understanding of how genetic and
570 environmental factors shape brain development. These insights may inform early
571 intervention strategies aimed at mitigating neurodevelopmental disparities, such as
572 targeted cognitive training, socioeconomic policy reforms, or school-based
573 enrichment programs that support children from disadvantaged backgrounds.

574 **Ethics Statement**

575 This study analyzed data from the Adolescent Brain Cognitive Development
576 (ABCD) study. All participants provided informed assent, and their parents or legal
577 guardians provided informed consent before participating in the ABCD study.

578

579 **Declaration of Competing Interest**

580 The authors declare no competing financial interests.

581

582

583 **Acknowledgments**

584 This work was supported by the National Research Foundation of Korea(NRF) grant

585 funded by the Korea government(MSIT) (No. 2021R1C1C1006503, RS-2023-

586 00266787, RS-2023-00265406, RS-2024-00421268), by Creative-Pioneering

587 Researchers Program through Seoul National University(No. 200-20240057), by

588 Semi-Supervised Learning Research Grant by SAMSUNG(No.A0426-20220118), by

589 Identify the network of brain preparation steps for concentration Research Grant by

590 LooxidLabs(No.339-20230001), by Institute of Information & communications

591 Technology Planning & Evaluation (IITP) grant funded by the Korea

592 government(MSIT) [NO.RS-2021-II211343, Artificial Intelligence Graduate School

593 Program (Seoul National University)] by the MSIT(Ministry of Science, ICT), Korea,

594 under the Global Research Support Program in the Digital Field program(RS-2024-

595 00421268) supervised by the IITP(Institute for Information & Communications

596 Technology Planning & Evaluation), by the National Supercomputing Center with

597 supercomputing resources including technical support(KSC-2023-CRE-0568) and by

598 the Ministry of Education of the Republic of Korea and the National Research

599 Foundation of Korea (NRF-2021S1A3A2A02090597), and by Artificial intelligence

600 industrial convergence cluster development project funded by the Ministry of Science

601 and ICT(MSIT, Korea) & Gwangju Metropolitan City.

602 **Data Availability**

603 All original data are publicly available from the NDA
604 (<https://nda.nih.gov/abcd/>). For rapid replication, we provide synthetic data, whose
605 distributions were matched to the original data, generated using conditional GAN for
606 tabular data (Xu et al., 2019) ([https://github.com/Transconnectome/ABCD-brain-](https://github.com/Transconnectome/ABCD-brain-network-SCCA)
607 [network-SCCA](https://github.com/Transconnectome/ABCD-brain-network-SCCA)).

608

609 **Code Availability**

610 Code is available from here: [https://github.com/Transconnectome/ABCD-](https://github.com/Transconnectome/ABCD-brain-network-SCCA)
611 [brain-network-SCCA](https://github.com/Transconnectome/ABCD-brain-network-SCCA).

612

613 **Author contributions**

- 614 ● Jungwoo Seo, Conceptualization, Data curation, Formal analysis, Investigation,
615 Methodology, Project administration, Software, Validation, Visualization, Writing –
616 original draft, Writing – review and editing
- 617 ● Eunji Lee, Data curation, Writing – original draft, Writing – review and editing
- 618 ● Bo-Gyeom Kim, Data curation, Writing – review and editing
- 619 ● Gakyung Kim, Data curation, Writing – original draft, Writing – review and editing
- 620 ● Yoonjung Yoonie Joo, Data curation, Writing – review and editing
- 621 ● Jiok Cha, Conceptualization, Supervision, Project administration, Funding
622 acquisition, Resources, Data curation, Writing – original draft, Writing – review and
623 editing.

625 **Tables**

626 **Table 1. Demographic information of the main analysis participants.**

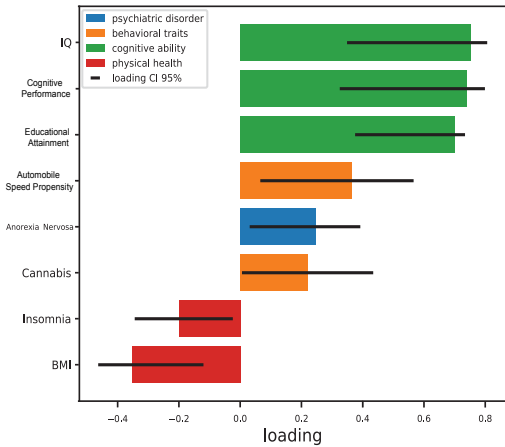
		PGS – BNM (n=5,784)		ENV – BNM (n=10,343)		Pheno – BNM (n=10,343)	
		train	test	train	test	train	test
N	total	4,626	1,157	8,274	2,069	8,274	2,069
Sex	Male	2,453	627	4,343	1,065	4,343	1,065
	Female	2,172	529	3,929	1,003	3,929	1,003
	Other	1	1	2	1	2	1
Race	White	3,225	802	4,350	1,128	4,350	1,128
	Black	29	7	1,194	295	1,194	295
	Hispanic	902	223	1,686	408	1,686	408
	Asian	8	0	173	37	173	37
	Other	462	125	871	201	871	201

627

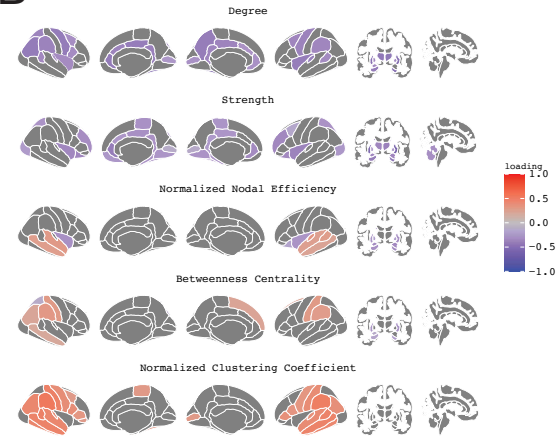
628 **Figure 1 Loading patterns of principal mode of sparse canonical correlation analysis.** The
629 SCCA results for the first modes of PGS-BNM, ENV-BNM, and Pheno-BNM. Only variables with 95%
630 confidence intervals, estimated from 5,000 bootstrap samples, that do not cross zero are shown.
631 Results for other variables can be found in supplementary figures. (A) The loadings of significant PGS
632 variables in the PGS-BNM mode 1. The error bars represent the 95% confidence interval of the
633 loading, estimated from the 5,000 bootstrap samples. The color of each bar represents the category
634 to which the variable belongs. (B) The loadings of significant nodal brain network measures in the
635 PGS-BNM mode 1. The loading patterns were visualized with R-package 'ggseg' (Mowinckel & Vidal-
636 Pineiro, 2020). (C) The loadings of significant environmental variables in the ENV-BNM mode 1. (D)
637 The loadings of significant nodal brain network measures in the ENV-BNM mode 1. (E) The loadings
638 of significant phenotype variables in the Pheno-BNM mode 1. (F) The loadings of significant nodal
639 brain network measures in the Pheno-BNM mode 1. (G) Brain network measures that their loadings
640 are commonly significant in PGS-BNM mode 1, ENV-BNM mode1, and Pheno-BNM mode1.

641

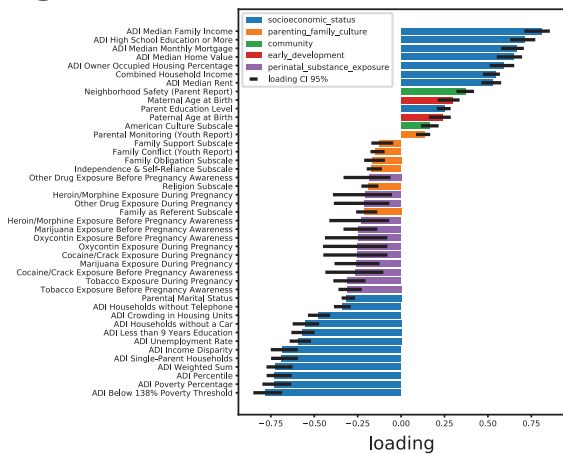
A Polygenic Scores Lodgings in PGS-BNM SCCA Mode1



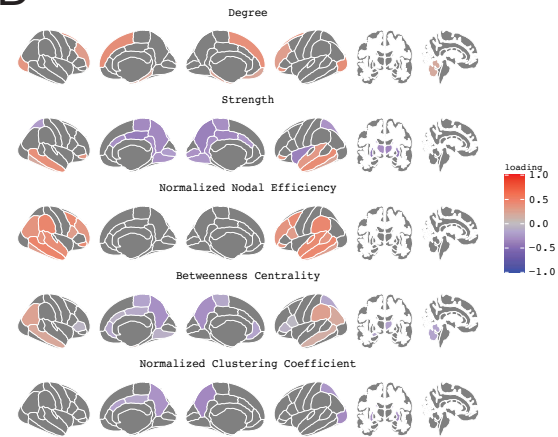
B Brain Network Lodgings in PGS-BNM SCCA Mode1



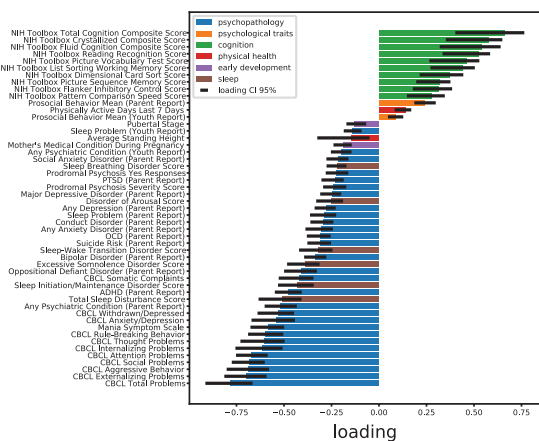
C Environmental Lodgings in ENV-BNM SCCA Mode1



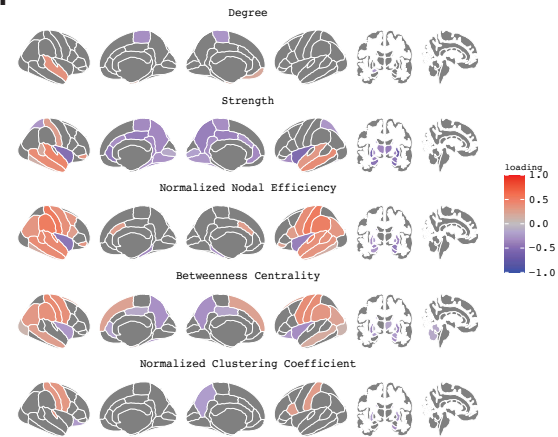
D Brain Network Lodgings in ENV-BNM SCCA Mode1



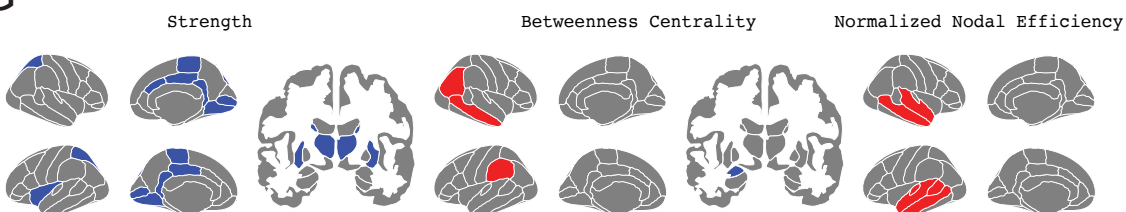
E Phenotype Lodgings in Pheno-BNM SCCA Mode1



F Brain Network Loadings in Pheno-BNM SCCA Mode1



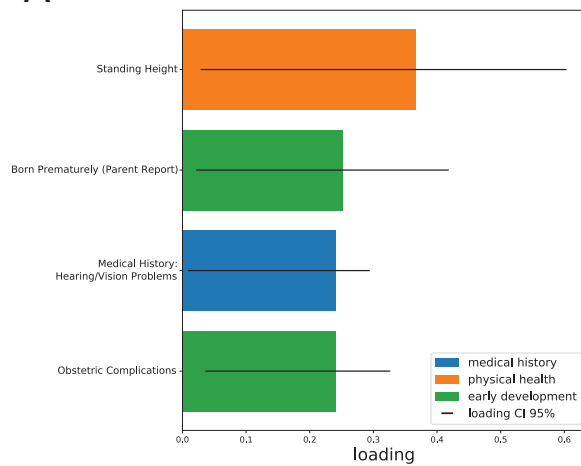
G Commonly Covarying Brain Network Measures



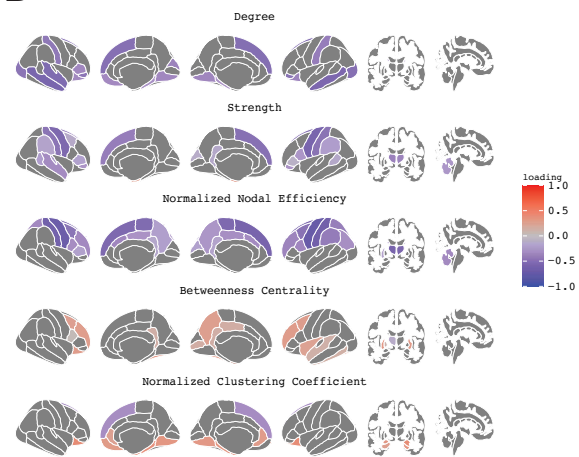
643

644 **Figure 2 Loading patterns of second mode of Pheno-BNM sparse canonical correlation**
645 **analysis.** (A) The loadings of significant phenotype variables in the Pheno-BNM SCCA mode 2. (B)
646 The loadings of significant nodal brain network measures in the Pheno-BNM SCCA mode 2.

A Phenotype Loadings in PGS-BNM SCCA Mode 2

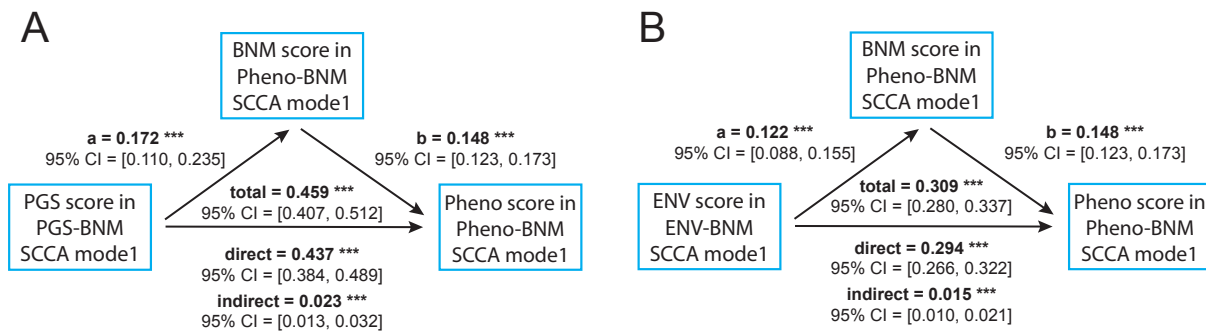


B Brain Network Loadings in PGS-BNM SCCA Mode 2



647

648 **Figure 3 Results of mediation analysis investigating the gene-brain network-phenotype and**
649 **environment-brain network-phenotype pathways.** Mediation analysis was conducted using
650 canonical variates derived from SCCA for polygenic scores, brain network measures, phenotypes,
651 and environmental factors. (A) Brain network properties mediating the relationship between polygenic
652 scores (reflecting cognitive ability) and phenotype scores (reflecting cognitive ability-
653 psychopathological traits). (B) Brain network properties mediating the relationship between
654 environmental factors (reflecting socioeconomic status) and phenotype scores (cognitive ability-
655 psychopathological traits).



656

657

658 References

- 659 Akiyama, M., Okada, Y., Kanai, M., Takahashi, A., Momozawa, Y., Ikeda, M., Iwata, N.,
660 Ikegawa, S., Hirata, M., Matsuda, K., Iwasaki, M., Yamaji, T., Sawada, N., Hachiya,
661 T., Tanno, K., Shimizu, A., Hozawa, A., Minegishi, N., Tsugane, S., . . . Kamatani, Y.
662 (2017). Genome-wide association study identifies 112 new loci for body mass index
663 in the Japanese population. *Nat Genet*, *49*(10), 1458-1467.
664 <https://doi.org/10.1038/ng.3951>
- 665 Alexander-Bloch, A. F., Gogtay, N., Meunier, D., Birn, R., Clasen, L., Lalonde, F., Lenroot, R.,
666 Giedd, J., & Bullmore, E. T. (2010). Disrupted modularity and local connectivity of
667 brain functional networks in childhood-onset schizophrenia. *Front Syst Neurosci*, *4*,
668 147. <https://doi.org/10.3389/fnsys.2010.00147>
- 669 Alnaes, D., Kaufmann, T., Marquand, A. F., Smith, S. M., & Westlye, L. T. (2020). Patterns of
670 sociocognitive stratification and perinatal risk in the child brain. *Proc Natl Acad Sci U*
671 *S A*, *117*(22), 12419-12427. <https://doi.org/10.1073/pnas.2001517117>
- 672 Bai, S., Liu, W., & Guan, Y. (2021). The Visuospatial and Sensorimotor Functions of
673 Posterior Parietal Cortex in Drawing Tasks: A Review. *Front Aging Neurosci*, *13*,
674 717002. <https://doi.org/10.3389/fnagi.2021.717002>
- 675 Baker, S. T., Lubman, D. I., Yucel, M., Allen, N. B., Whittle, S., Fulcher, B. D., Zalesky, A., &
676 Fornito, A. (2015). Developmental Changes in Brain Network Hub Connectivity in
677 Late Adolescence. *J Neurosci*, *35*(24), 9078-9087.
678 <https://doi.org/10.1523/JNEUROSCI.5043-14.2015>
- 679 Basten, U., Hilger, K., & Fiebach, C. J. (2015). Where smart brains are different: A
680 quantitative meta-analysis of functional and structural brain imaging studies on
681 intelligence. *Intelligence*, *51*, 10-27.
- 682 Bathelt, J., Gathercole, S. E., Butterfield, S., team, C., & Astle, D. E. (2018). Children's
683 academic attainment is linked to the global organization of the white matter
684 connectome. *Dev Sci*, *21*(5), e12662. <https://doi.org/10.1111/desc.12662>
- 685 Bathelt, J., Gathercole, S. E., Johnson, A., & Astle, D. E. (2018). Differences in brain
686 morphology and working memory capacity across childhood. *Dev Sci*, *21*(3), e12579.
687 <https://doi.org/10.1111/desc.12579>
- 688 Bethlehem, R. A. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C.,
689 Adler, S., Alexopoulos, G. S., Anagnostou, E., Areces-Gonzalez, A., Astle, D. E.,
690 Auyeung, B., Ayub, M., Bae, J., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S. A.,
691 Benegal, V., . . . Alexander-Bloch, A. F. (2022). Brain charts for the human lifespan.
692 *Nature*, *604*(7906), 525-533. <https://doi.org/10.1038/s41586-022-04554-y>
- 693 Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic
694 system? A critical review and meta-analysis of 120 functional neuroimaging studies.
695 *Cereb Cortex*, *19*(12), 2767-2796. <https://doi.org/10.1093/cercor/bhp055>
- 696 Bipolar, D., Schizophrenia Working Group of the Psychiatric Genomics Consortium.
697 Electronic address, d. r. v. e., Bipolar, D., & Schizophrenia Working Group of the
698 Psychiatric Genomics, C. (2018). Genomic Dissection of Bipolar Disorder and
699 Schizophrenia, Including 28 Subphenotypes. *Cell*, *173*(7), 1705-1715 e1716.
700 <https://doi.org/10.1016/j.cell.2018.05.046>
- 701 Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network:
702 anatomy, function, and relevance to disease. *Ann N Y Acad Sci*, *1124*, 1-38.
703 <https://doi.org/10.1196/annals.1440.011>
- 704 Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and
705 cognitive control. *Curr Opin Neurobiol*, *17*(2), 243-250.
706 <https://doi.org/10.1016/j.conb.2007.02.005>
- 707 Calamante, F., Tournier, J. D., Jackson, G. D., & Connelly, A. (2010). Track-density imaging
708 (TDI): super-resolution white matter imaging using whole-brain track-density

- 709 mapping. *Neuroimage*, 53(4), 1233-1243.
710 <https://doi.org/10.1016/j.neuroimage.2010.07.024>
- 711 Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M.,
712 Soules, M. E., Teslovich, T., Dellarco, D. V., Garavan, H., Orr, C. A., Wager, T. D.,
713 Banich, M. T., Speer, N. K., Sutherland, M. T., Riedel, M. C., Dick, A. S., Bjork, J. M.,
714 Thomas, K. M., . . . Workgroup, A. I. A. (2018). The Adolescent Brain Cognitive
715 Development (ABCD) study: Imaging acquisition across 21 sites. *Dev Cogn*
716 *Neurosci*, 32, 43-54. <https://doi.org/10.1016/j.dcn.2018.03.001>
- 717 Collin, G., Scholtens, L. H., Kahn, R. S., Hillegers, M. H. J., & van den Heuvel, M. P. (2017).
718 Affected Anatomical Rich Club and Structural-Functional Coupling in Young Offspring
719 of Schizophrenia and Bipolar Disorder Patients. *Biol Psychiatry*, 82(10), 746-755.
720 <https://doi.org/10.1016/j.biopsych.2017.06.013>
- 721 Conomos, M. P., Miller, M. B., & Thornton, T. A. (2015). Robust inference of population
722 structure for ancestry prediction and correction of stratification in the presence of
723 relatedness. *Genet Epidemiol*, 39(4), 276-293. <https://doi.org/10.1002/gepi.21896>
- 724 Conomos, M. P., Reiner, A. P., Weir, B. S., & Thornton, T. A. (2016). Model-free Estimation of
725 Recent Genetic Relatedness. *Am J Hum Genet*, 98(1), 127-148.
726 <https://doi.org/10.1016/j.ajhg.2015.11.022>
- 727 Cross-Disorder Group of the Psychiatric Genomics, C. (2013). Identification of risk loci with
728 shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*,
729 381(9875), 1371-1379. [https://doi.org/10.1016/S0140-6736\(12\)62129-1](https://doi.org/10.1016/S0140-6736(12)62129-1)
- 730 Culham, J. C., & Kanwisher, N. G. (2001). Neuroimaging of cognitive functions in human
731 parietal cortex. *Curr Opin Neurobiol*, 11(2), 157-163. [https://doi.org/10.1016/s0959-4388\(00\)00191-4](https://doi.org/10.1016/s0959-4388(00)00191-4)
- 732
- 733 Das, S., Forer, L., Schonherr, S., Sidore, C., Locke, A. E., Kwong, A., Vrieze, S. I., Chew, E.
734 Y., Levy, S., McGue, M., Schlessinger, D., Stambolian, D., Loh, P. R., Iacono, W. G.,
735 Swaroop, A., Scott, L. J., Cucca, F., Kronenberg, F., Boehnke, M., . . . Fuchsberger,
736 C. (2016). Next-generation genotype imputation service and methods. *Nat Genet*,
737 48(10), 1284-1287. <https://doi.org/10.1038/ng.3656>
- 738 Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., Baldursson,
739 G., Belliveau, R., Bybjerg-Grauholm, J., Baekvad-Hansen, M., Cerrato, F., Chambert,
740 K., Churchhouse, C., Dumont, A., Eriksson, N., Gandal, M., Goldstein, J. I., Grasby,
741 K. L., Grove, J., . . . Neale, B. M. (2019). Discovery of the first genome-wide
742 significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet*, 51(1), 63-75.
743 <https://doi.org/10.1038/s41588-018-0269-7>
- 744 Deschamps, I., Baum, S. R., & Gracco, V. L. (2014). On the role of the supramarginal gyrus
745 in phonological processing and verbal working memory: evidence from rTMS studies.
746 *Neuropsychologia*, 53, 39-46. <https://doi.org/10.1016/j.neuropsychologia.2013.10.015>
- 747 Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner,
748 R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006).
749 An automated labeling system for subdividing the human cerebral cortex on MRI
750 scans into gyral based regions of interest. *Neuroimage*, 31(3), 968-980.
751 <https://doi.org/10.1016/j.neuroimage.2006.01.021>
- 752 du Boisgueheneuc, F., Levy, R., Volle, E., Seassau, M., Duffau, H., Kinkingnehun, S.,
753 Samson, Y., Zhang, S., & Dubois, B. (2006). Functions of the left superior frontal
754 gyrus in humans: a lesion study. *Brain*, 129(Pt 12), 3315-3328.
755 <https://doi.org/10.1093/brain/awl244>
- 756 El Marroun, H., Tiemeier, H., Franken, I. H., Jaddoe, V. W., van der Lugt, A., Verhulst, F. C.,
757 Lahey, B. B., & White, T. (2016). Prenatal Cannabis and Tobacco Exposure in
758 Relation to Brain Morphology: A Prospective Neuroimaging Study in Young Children.
759 *Biol Psychiatry*, 79(12), 971-979. <https://doi.org/10.1016/j.biopsych.2015.08.024>
- 760 Fernandez-Cabello, S., Alnaes, D., van der Meer, D., Dahl, A., Holm, M., Kjelkenes, R.,
761 Maximov, I., Norbom, L. B., Pedersen, M. L., Voldsbekk, I., Andreassen, O. A., &
762 Westlye, L. T. (2022). Associations between brain imaging and polygenic scores of

- 763 mental health and educational attainment in children aged 9-11. *Neuroimage*, 263,
764 119611. <https://doi.org/10.1016/j.neuroimage.2022.119611>
- 765 Fornito, A., Zalesky, A., & Breakspear, M. (2015). The connectomics of brain disorders. *Nat*
766 *Rev Neurosci*, 16(3), 159-172. <https://doi.org/10.1038/nrn3901>
- 767 Ge, T., Chen, C. Y., Ni, Y., Feng, Y. A., & Smoller, J. W. (2019). Polygenic prediction via
768 Bayesian regression and continuous shrinkage priors. *Nat Commun*, 10(1), 1776.
769 <https://doi.org/10.1038/s41467-019-09718-5>
- 770 Genomes Project, C., Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M.,
771 Korb, J. O., Marchini, J. L., McCarthy, S., McVean, G. A., & Abecasis, G. R. (2015).
772 A global reference for human genetic variation. *Nature*, 526(7571), 68-74.
773 <https://doi.org/10.1038/nature15393>
- 774 Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., Pallesen, J., Agerbo,
775 E., Andreassen, O. A., Anney, R., Awasthi, S., Belliveau, R., Bettella, F., Buxbaum, J.
776 D., Bybjerg-Grauholm, J., Baekvad-Hansen, M., Cerrato, F., Chambert, K.,
777 Christensen, J. H., . . . Borglum, A. D. (2019). Identification of common genetic risk
778 variants for autism spectrum disorder. *Nat Genet*, 51(3), 431-444.
779 <https://doi.org/10.1038/s41588-019-0344-8>
- 780 Gu, Z., Barch, D. M., & Luo, Q. (2024). Prenatal substance exposure and child health:
781 Understanding the role of environmental factors, genetics, and brain development.
782 *PNAS Nexus*, 3(1), pgae003. <https://doi.org/10.1093/pnasnexus/pgae003>
- 783 Heller, A. S., Cohen, A. O., Dreyfuss, M. F., & Casey, B. J. (2016). Changes in cortico-
784 subcortical and subcortico-subcortical connectivity impact cognitive control to
785 emotional cues across development. *Soc Cogn Affect Neurosci*, 11(12), 1910-1918.
786 <https://doi.org/10.1093/scan/nsw097>
- 787 Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28, 321-377.
788 <https://doi.org/10.1093/biomet/28.3-4.321>
- 789 Howard, D. M., Adams, M. J., Clarke, T. K., Hafferty, J. D., Gibson, J., Shiri, M., Coleman,
790 J. R. I., Hagenaars, S. P., Ward, J., Wigmore, E. M., Alloza, C., Shen, X., Barbu, M.
791 C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I.
792 J., . . . McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies
793 102 independent variants and highlights the importance of the prefrontal brain
794 regions. *Nat Neurosci*, 22(3), 343-352. <https://doi.org/10.1038/s41593-018-0326-7>
- 795 Hwang, K., Hallquist, M. N., & Luna, B. (2013). The development of hub architecture in the
796 human functional brain network. *Cereb Cortex*, 23(10), 2380-2393.
797 <https://doi.org/10.1093/cercor/bhs227>
- 798 International Obsessive Compulsive Disorder Foundation Genetics, C., & Studies, O. C. D.
799 C. G. A. (2018). Revealing the complex genetic architecture of obsessive-compulsive
800 disorder using meta-analysis. *Mol Psychiatry*, 23(5), 1181-1188.
801 <https://doi.org/10.1038/mp.2017.154>
- 802 Jansen, P. R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A. R., de
803 Leeuw, C. A., Benjamins, J. S., Munoz-Manchado, A. B., Nagel, M., Savage, J. E.,
804 Tiemeier, H., White, T., and Me Research, T., Tung, J. Y., Hinds, D. A., Vacic, V.,
805 Wang, X., Sullivan, P. F., . . . Posthuma, D. (2019). Genome-wide analysis of
806 insomnia in 1,331,010 individuals identifies new risk loci and functional pathways.
807 *Nat Genet*, 51(3), 394-403. <https://doi.org/10.1038/s41588-018-0333-3>
- 808 Jernigan, T. L., Brown, S. A., & Dowling, G. J. (2018). The Adolescent Brain Cognitive
809 Development Study. *J Res Adolesc*, 28(1), 154-156.
810 <https://doi.org/10.1111/jora.12374>
- 811 Jeurissen, B., Descoteaux, M., Mori, S., & Leemans, A. (2019). Diffusion MRI fiber
812 tractography of the brain. *NMR Biomed*, 32(4), e3785.
813 <https://doi.org/10.1002/nbm.3785>
- 814 Joo, Y. Y., Lee, E., Kim, B. G., Kim, G., Seo, J., & Cha, J. (2024). Polygenic architecture of
815 brain structure and function, behaviors, and psychopathologies in children. *bioRxiv*.
816 <https://doi.org/10.1101/2024.05.22.595444>

- 817 Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of
818 intelligence: converging neuroimaging evidence. *Behav Brain Sci*, 30(2), 135-154;
819 discussion 154-187. <https://doi.org/10.1017/S0140525X07001185>
- 820 Karlsson Linner, R., Biroli, P., Kong, E., Meddens, S. F. W., Wedow, R., Fontana, M. A.,
821 Lebreton, M., Tino, S. P., Abdellaoui, A., Hammerschlag, A. R., Nivard, M. G., Okbay,
822 A., Rietveld, C. A., Timshel, P. N., Trzaskowski, M., Vlaming, R., Zund, C. L., Bao, Y.,
823 Buzdugan, L., . . . Beauchamp, J. P. (2019). Genome-wide association analyses of
824 risk tolerance and risky behaviors in over 1 million individuals identify hundreds of
825 loci and shared genetic influences. *Nat Genet*, 51(2), 245-257.
826 <https://doi.org/10.1038/s41588-018-0309-3>
- 827 Kim, D. J., Davis, E. P., Sandman, C. A., Glynn, L., Sporns, O., O'Donnell, B. F., & Hetrick,
828 W. P. (2019). Childhood poverty and the organization of structural brain connectome.
829 *Neuroimage*, 184, 409-416. <https://doi.org/10.1016/j.neuroimage.2018.09.041>
- 830 Kim, D. J., Davis, E. P., Sandman, C. A., Sporns, O., O'Donnell, B. F., Buss, C., & Hetrick, W.
831 P. (2016). Children's intellectual ability is associated with structural network integrity.
832 *Neuroimage*, 124(Pt A), 550-556. <https://doi.org/10.1016/j.neuroimage.2015.09.012>
- 833 Kim, K., Joo, Y. Y., Ahn, G., Wang, H. H., Moon, S. Y., Kim, H., Ahn, W. Y., & Cha, J. (2022).
834 The sexual brain, genes, and cognition: A machine-predicted brain sex score
835 explains individual differences in cognitive intelligence and genetic influence in young
836 children. *Hum Brain Mapp*, 43(12), 3857-3872. <https://doi.org/10.1002/hbm.25888>
- 837 Koenis, M. M., Brouwer, R. M., van den Heuvel, M. P., Mandl, R. C., van Soelen, I. L., Kahn,
838 R. S., Boomsma, D. I., & Hulshoff Pol, H. E. (2015). Development of the brain's
839 structural network efficiency in early adolescence: A longitudinal DTI twin study. *Hum*
840 *Brain Mapp*, 36(12), 4938-4953. <https://doi.org/10.1002/hbm.22988>
- 841 Lam, M., Chen, C. Y., Li, Z., Martin, A. R., Bryois, J., Ma, X., Gaspar, H., Ikeda, M.,
842 Benyamin, B., Brown, B. C., Liu, R., Zhou, W., Guan, L., Kamatani, Y., Kim, S. W.,
843 Kubo, M., Kusumawardhani, A., Liu, C. M., Ma, H., . . . Huang, H. (2019).
844 Comparative genetic architectures of schizophrenia in East Asian and European
845 populations. *Nat Genet*, 51(12), 1670-1678. [https://doi.org/10.1038/s41588-019-](https://doi.org/10.1038/s41588-019-0512-x)
846 [0512-x](https://doi.org/10.1038/s41588-019-0512-x)
- 847 Langen, C. D., Muetzel, R., Blanken, L., van der Lugt, A., Tiemeier, H., Verhulst, F., Niessen,
848 W. J., & White, T. (2018). Differential patterns of age-related cortical and subcortical
849 functional connectivity in 6-to-10 year old children: A connectome-wide association
850 study. *Brain Behav*, 8(8), e01031. <https://doi.org/10.1002/brb3.1031>
- 851 Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghziyan, O., Zacher, M., Nguyen-Viet, T. A.,
852 Bowers, P., Sidorenko, J., Karlsson Linner, R., Fontana, M. A., Kundu, T., Lee, C., Li,
853 H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., . . . Cesarini, D.
854 (2018). Gene discovery and polygenic prediction from a genome-wide association
855 study of educational attainment in 1.1 million individuals. *Nat Genet*, 50(8), 1112-
856 1121. <https://doi.org/10.1038/s41588-018-0147-3>
- 857 Lehericy, S., Bardinnet, E., Tremblay, L., Van de Moortele, P. F., Pochon, J. B., Dormont, D.,
858 Kim, D. S., Yelnik, J., & Ugurbil, K. (2006). Motor control in basal ganglia circuits
859 using fMRI and brain atlas approaches. *Cereb Cortex*, 16(2), 149-161.
860 <https://doi.org/10.1093/cercor/bhi089>
- 861 Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., Powell, C.,
862 Vedantam, S., Buchkovich, M. L., Yang, J., Croteau-Chonka, D. C., Esko, T., Fall, T.,
863 Ferreira, T., Gustafsson, S., Kutalik, Z., Luan, J., Magi, R., Randall, J. C., . . .
864 Speliotes, E. K. (2015). Genetic studies of body mass index yield new insights for
865 obesity biology. *Nature*, 518(7538), 197-206. <https://doi.org/10.1038/nature14177>
- 866 Loh, P. R., Danecek, P., Palamara, P. F., Fuchsberger, C., Y, A. R., H, K. F., Schoenherr, S.,
867 Forer, L., McCarthy, S., Abecasis, G. R., Durbin, R., & A, L. P. (2016). Reference-
868 based phasing using the Haplotype Reference Consortium panel. *Nat Genet*, 48(11),
869 1443-1448. <https://doi.org/10.1038/ng.3679>
- 870 Luna, B., Padmanabhan, A., & O'Hearn, K. (2010). What has fMRI told us about the

- 871 development of cognitive control through adolescence? *Brain Cogn*, 72(1), 101-113.
872 <https://doi.org/10.1016/j.bandc.2009.08.005>
- 873 Ma, J., Kang, H. J., Kim, J. Y., Jeong, H. S., Im, J. J., Namgung, E., Kim, M. J., Lee, S., Kim,
874 T. D., Oh, J. K., Chung, Y. A., Lyoo, I. K., Lim, S. M., & Yoon, S. (2017). Network
875 attributes underlying intellectual giftedness in the developing brain. *Sci Rep*, 7(1),
876 11321. <https://doi.org/10.1038/s41598-017-11593-3>
- 877 Menon, V. (2013). Developmental pathways to functional brain networks: emerging
878 principles. *Trends Cogn Sci*, 17(12), 627-640.
879 <https://doi.org/10.1016/j.tics.2013.09.015>
- 880 Metz, V. E., Brown, Q. L., Martins, S. S., & Palamar, J. J. (2018). Characteristics of drug use
881 among pregnant women in the United States: Opioid and non-opioid illegal drug use.
882 *Drug Alcohol Depend*, 183, 261-266.
883 <https://doi.org/10.1016/j.drugalcdep.2017.11.010>
- 884 Mistic, B., Betzel, R. F., de Reus, M. A., van den Heuvel, M. P., Berman, M. G., McIntosh, A.
885 R., & Sporns, O. (2016). Network-Level Structure-Function Relationships in Human
886 Neocortex. *Cereb Cortex*, 26(7), 3285-3296. <https://doi.org/10.1093/cercor/bhw089>
- 887 Modabbernia, A., Janiri, D., Doucet, G. E., Reichenberg, A., & Frangou, S. (2021).
888 Multivariate Patterns of Brain-Behavior-Environment Associations in the Adolescent
889 Brain and Cognitive Development Study. *Biol Psychiatry*, 89(5), 510-520.
890 <https://doi.org/10.1016/j.biopsych.2020.08.014>
- 891 Mowinckel, A. M., & Vidal-Pineiro, D. (2020). Visualization of Brain Statistics With R
892 Packages ggseg and ggseg3d. *Advances in Methods and Practices in Psychological*
893 *Science*, 3(4), 466-483. <https://doi.org/10.1177/2515245920928009>
- 894 Mravcik, V., Nechanska, B., Gabrhelik, R., Handal, M., Mahic, M., & Skurtveit, S. (2020).
895 Socioeconomic characteristics of women with substance use disorder during
896 pregnancy and neonatal outcomes in their newborns: A national registry study from
897 the Czech Republic. *Drug Alcohol Depend*, 209, 107933.
898 <https://doi.org/10.1016/j.drugalcdep.2020.107933>
- 899 Nagel, M., Jansen, P. R., Stringer, S., Watanabe, K., de Leeuw, C. A., Bryois, J., Savage, J.
900 E., Hammerschlag, A. R., Skene, N. G., Munoz-Manchado, A. B., andMe Research,
901 T., White, T., Tiemeier, H., Linnarsson, S., Hjerling-Leffler, J., Polderman, T. J. C.,
902 Sullivan, P. F., van der Sluis, S., & Posthuma, D. (2018). Meta-analysis of genome-
903 wide association studies for neuroticism in 449,484 individuals identifies novel
904 genetic loci and pathways. *Nat Genet*, 50(7), 920-927.
905 <https://doi.org/10.1038/s41588-018-0151-7>
- 906 Namkung, H., Kim, S. H., & Sawa, A. (2017). The Insula: An Underestimated Brain Area in
907 Clinical Neuroscience, Psychiatry, and Neurology. *Trends Neurosci*, 40(4), 200-207.
908 <https://doi.org/10.1016/j.tins.2017.02.002>
- 909 Nievergelt, C. M., Maihofer, A. X., Klengel, T., Atkinson, E. G., Chen, C. Y., Choi, K. W.,
910 Coleman, J. R. I., Dalvie, S., Duncan, L. E., Gelernter, J., Levey, D. F., Logue, M. W.,
911 Polimanti, R., Provost, A. C., Ratanatharathorn, A., Stein, M. B., Torres, K., Aiello, A.
912 E., Almli, L. M., . . . Koenen, K. C. (2019). International meta-analysis of PTSD
913 genome-wide association studies identifies sex- and ancestry-specific genetic risk
914 loci. *Nat Commun*, 10(1), 4558. <https://doi.org/10.1038/s41467-019-12576-w>
- 915 Okbay, A., Baselmans, B. M., De Neve, J. E., Turley, P., Nivard, M. G., Fontana, M. A.,
916 Meddens, S. F., Linner, R. K., Rietveld, C. A., Derringer, J., Gratten, J., Lee, J. J., Liu,
917 J. Z., de Vlaming, R., Ahluwalia, T. S., Buchwald, J., Cavadino, A., Frazier-Wood, A.
918 C., Furlotte, N. A., . . . Cesarini, D. (2016). Genetic variants associated with
919 subjective well-being, depressive symptoms, and neuroticism identified through
920 genome-wide analyses. *Nat Genet*, 48(6), 624-633. <https://doi.org/10.1038/ng.3552>
- 921 Oldham, S., & Fornito, A. (2019). The development of brain network hubs. *Dev Cogn*
922 *Neurosci*, 36, 100607. <https://doi.org/10.1016/j.dcn.2018.12.005>
- 923 Otowa, T., Hek, K., Lee, M., Byrne, E. M., Mirza, S. S., Nivard, M. G., Bigdeli, T., Aggen, S.
924 H., Adkins, D., Wolen, A., Fanous, A., Keller, M. C., Castelao, E., Kutalik, Z., der

- 925 Auwera, S. V., Homuth, G., Nauck, M., Teumer, A., Milaneschi, Y., . . . Hettrema, J. M.
926 (2016). Meta-analysis of genome-wide association studies of anxiety disorders. *Mol*
927 *Psychiatry*, 21(10), 1485. <https://doi.org/10.1038/mp.2016.11>
- 928 Pasman, J. A., Verweij, K. J. H., Gerring, Z., Stringer, S., Sanchez-Roige, S., Treur, J. L.,
929 Abdellaoui, A., Nivard, M. G., Baselmans, B. M. L., Ong, J. S., Ip, H. F., van der Zee,
930 M. D., Bartels, M., Day, F. R., Fontanillas, P., Elson, S. L., and Me Research, T., de
931 Wit, H., Davis, L. K., . . . Vink, J. M. (2019). Author Correction: GWAS of lifetime
932 cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a
933 causal effect of schizophrenia liability. *Nat Neurosci*, 22(7), 1196.
934 <https://doi.org/10.1038/s41593-019-0402-7>
- 935 Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge
936 during adolescence? *Nat Rev Neurosci*, 9(12), 947-957.
937 <https://doi.org/10.1038/nrn2513>
- 938 Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of
939 heard speech, spoken language and reading. *Neuroimage*, 62(2), 816-847.
940 <https://doi.org/10.1016/j.neuroimage.2012.04.062>
- 941 Rakesh, D., Zalesky, A., & Whittle, S. (2022). Assessment of Parent Income and Education,
942 Neighborhood Disadvantage, and Child Brain Structure. *JAMA Netw Open*, 5(8),
943 e2226208. <https://doi.org/10.1001/jamanetworkopen.2022.26208>
- 944 Ross, E. J., Graham, D. L., Money, K. M., & Stanwood, G. D. (2015). Developmental
945 consequences of fetal exposure to drugs: what we know and what we still must learn.
946 *Neuropsychopharmacology*, 40(1), 61-87. <https://doi.org/10.1038/npp.2014.147>
- 947 Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses
948 and interpretations. *Neuroimage*, 52(3), 1059-1069.
949 <https://doi.org/10.1016/j.neuroimage.2009.10.003>
- 950 Rudie, J. D., Brown, J. A., Beck-Pancer, D., Hernandez, L. M., Dennis, E. L., Thompson, P.
951 M., Bookheimer, S. Y., & Dapretto, M. (2012). Altered functional and structural brain
952 network organization in autism. *Neuroimage Clin*, 2, 79-94.
953 <https://doi.org/10.1016/j.nicl.2012.11.006>
- 954 Sato, J. R., Salum, G. A., Gadelha, A., Vieira, G., Zugman, A., Picon, F. A., Pan, P. M.,
955 Hoexter, M. Q., Anes, M., Moura, L. M., Del'Aquila, M. A., Crossley, N., Amaro Junior,
956 E., McGuire, P., Lacerda, A. L., Rohde, L. A., Miguel, E. C., Jackowski, A. P., &
957 Bressan, R. A. (2015). Decreased centrality of subcortical regions during the
958 transition to adolescence: a functional connectivity study. *Neuroimage*, 104, 44-51.
959 <https://doi.org/10.1016/j.neuroimage.2014.09.063>
- 960 Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., Nagel,
961 M., Awasthi, S., Barr, P. B., Coleman, J. R. I., Grasby, K. L., Hammerschlag, A. R.,
962 Kaminski, J. A., Karlsson, R., Krapohl, E., Lam, M., Nygaard, M., Reynolds, C. A.,
963 Trampush, J. W., . . . Posthuma, D. (2018). Genome-wide association meta-analysis
964 in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat*
965 *Genet*, 50(7), 912-919. <https://doi.org/10.1038/s41588-018-0152-6>
- 966 Shen, H., Gelaye, B., Huang, H., Rondon, M. B., Sanchez, S., & Duncan, L. E. (2020).
967 Polygenic prediction and GWAS of depression, PTSD, and suicidal ideation/self-
968 harm in a Peruvian cohort. *Neuropsychopharmacology*, 45(10), 1595-1602.
969 <https://doi.org/10.1038/s41386-020-0603-5>
- 970 Sherman, S. M., & Guillery, R. W. (2002). The role of the thalamus in the flow of information
971 to the cortex. *Philos Trans R Soc Lond B Biol Sci*, 357(1428), 1695-1708.
972 <https://doi.org/10.1098/rstb.2002.1161>
- 973 Smith, R. E., Tournier, J. D., Calamante, F., & Connelly, A. (2013). SIFT: Spherical-
974 deconvolution informed filtering of tractograms. *Neuroimage*, 67, 298-312.
975 <https://doi.org/10.1016/j.neuroimage.2012.11.049>
- 976 Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E., Glasser, M. F.,
977 Ugurbil, K., Barch, D. M., Van Essen, D. C., & Miller, K. L. (2015). A positive-negative
978 mode of population covariation links brain connectivity, demographics and behavior.

- 979 *Nat Neurosci*, 18(11), 1565-1567. <https://doi.org/10.1038/nn.4125>
- 980 Sotiropoulos, S. N., & Zalesky, A. (2019). Building connectomes using diffusion MRI: why,
981 how and but. *NMR Biomed*, 32(4), e3752. <https://doi.org/10.1002/nbm.3752>
- 982 Stahl, E. A., Breen, G., Forstner, A. J., McQuillin, A., Ripke, S., Trubetskov, V., Mattheisen,
983 M., Wang, Y., Coleman, J. R. I., Gaspar, H. A., de Leeuw, C. A., Steinberg, S.,
984 Pavlides, J. M. W., Trzaskowski, M., Byrne, E. M., Pers, T. H., Holmans, P. A.,
985 Richards, A. L., Abbott, L., . . . Bipolar Disorder Working Group of the Psychiatric
986 Genomics, C. (2019). Genome-wide association study identifies 30 loci associated
987 with bipolar disorder. *Nat Genet*, 51(5), 793-803. <https://doi.org/10.1038/s41588-019-0397-8>
- 988
- 989 Suprano, I., Kocevar, G., Stamile, C., Hannoun, S., Fournier, P., Revol, O., Nusbaum, F., &
990 Sappey-Marini, D. (2020). White matter microarchitecture and structural network
991 integrity correlate with children intelligence quotient. *Sci Rep*, 10(1), 20722.
992 <https://doi.org/10.1038/s41598-020-76528-x>
- 993 Thompson, B. L., Levitt, P., & Stanwood, G. D. (2009). Prenatal exposure to drugs: effects
994 on brain development and implications for policy and education. *Nat Rev Neurosci*,
995 10(4), 303-312. <https://doi.org/10.1038/nrn2598>
- 996 Tooley, U. A., Mackey, A. P., Ciric, R., Ruparel, K., Moore, T. M., Gur, R. C., Gur, R. E.,
997 Satterthwaite, T. D., & Bassett, D. S. (2020). Associations between Neighborhood
998 SES and Functional Brain Network Development. *Cereb Cortex*, 30(1), 1-19.
999 <https://doi.org/10.1093/cercor/bhz066>
- 1000 Tournier, J. D., Calamante, F., & Connelly, A. (2007). Robust determination of the fibre
1001 orientation distribution in diffusion MRI: non-negativity constrained super-resolved
1002 spherical deconvolution. *Neuroimage*, 35(4), 1459-1472.
1003 <https://doi.org/10.1016/j.neuroimage.2007.02.016>
- 1004 Tournier, J. D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens,
1005 D., Jeurissen, B., Yeh, C. H., & Connelly, A. (2019). MRtrix3: A fast, flexible and open
1006 software framework for medical image processing and visualisation. *Neuroimage*,
1007 202, 116137. <https://doi.org/10.1016/j.neuroimage.2019.116137>
- 1008 Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained
1009 equations in R. *Journal of statistical software*, 45, 1-67.
- 1010 van den Heuvel, M. P., & Sporns, O. (2011). Rich-club organization of the human
1011 connectome. *J Neurosci*, 31(44), 15775-15786.
1012 <https://doi.org/10.1523/JNEUROSCI.3539-11.2011>
- 1013 van den Heuvel, M. P., van Soelen, I. L., Stam, C. J., Kahn, R. S., Boomsma, D. I., &
1014 Hulshoff Pol, H. E. (2013). Genetic control of functional brain network efficiency in
1015 children. *Eur Neuropsychopharmacol*, 23(1), 19-23.
1016 <https://doi.org/10.1016/j.euroneuro.2012.06.007>
- 1017 Visser, M., Jefferies, E., Embleton, K. V., & Lambon Ralph, M. A. (2012). Both the middle
1018 temporal gyrus and the ventral anterior temporal area are crucial for multimodal
1019 semantic processing: distortion-corrected fMRI evidence for a double gradient of
1020 information convergence in the temporal lobes. *J Cogn Neurosci*, 24(8), 1766-1778.
1021 https://doi.org/10.1162/jocn_a_00244
- 1022 Walters, R. K., Polimanti, R., Johnson, E. C., McClintick, J. N., Adams, M. J., Adkins, A. E.,
1023 Aliev, F., Bacanu, S. A., Batzler, A., Bertelsen, S., Biernacka, J. M., Bigdeli, T. B.,
1024 Chen, L. S., Clarke, T. K., Chou, Y. L., Degenhardt, F., Docherty, A. R., Edwards, A.
1025 C., Fontanillas, P., . . . Agrawal, A. (2018). Transancestral GWAS of alcohol
1026 dependence reveals common genetic underpinnings with psychiatric disorders. *Nat*
1027 *Neurosci*, 21(12), 1656-1669. <https://doi.org/10.1038/s41593-018-0275-1>
- 1028 Wang, H. T., Smallwood, J., Mourao-Miranda, J., Xia, C. H., Satterthwaite, T. D., Bassett, D.
1029 S., & Bzdok, D. (2020). Finding the needle in a high-dimensional haystack: Canonical
1030 correlation analysis for neuroscientists. *Neuroimage*, 216, 116745.
1031 <https://doi.org/10.1016/j.neuroimage.2020.116745>
- 1032 Watson, H. J., Yilmaz, Z., Thornton, L. M., Hubel, C., Coleman, J. R. I., Gaspar, H. A.,

- 1033 Bryois, J., Hinney, A., Leppa, V. M., Mattheisen, M., Medland, S. E., Ripke, S., Yao,
1034 S., Giusti-Rodriguez, P., Anorexia Nervosa Genetics, I., Hanscombe, K. B., Purves,
1035 K. L., Eating Disorders Working Group of the Psychiatric Genomics, C., Adan, R. A.
1036 H., . . . Bulik, C. M. (2019). Genome-wide association study identifies eight risk loci
1037 and implicates metabo-psychiatric origins for anorexia nervosa. *Nat Genet*, *51*(8),
1038 1207-1214. <https://doi.org/10.1038/s41588-019-0439-2>
- 1039 Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J.,
1040 Carlozzi, N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., Fox, N. A., Beaumont, J. L.,
1041 Mungas, D., Nowinski, C. J., Richler, J., Deocampo, J. A., Anderson, J. E., Manly, J.
1042 J., Borosh, B., . . . Gershon, R. C. (2013). Cognition assessment using the NIH
1043 Toolbox. *Neurology*, *80*(11 Suppl 3), S54-64.
1044 <https://doi.org/10.1212/WNL.0b013e3182872ded>
- 1045 Witten, D. M., Tibshirani, R., & Hastie, T. (2009). A penalized matrix decomposition, with
1046 applications to sparse principal components and canonical correlation analysis.
1047 *Biostatistics*, *10*(3), 515-534. <https://doi.org/10.1093/biostatistics/kxp008>
- 1048 Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A.,
1049 Adams, M. J., Agerbo, E., Air, T. M., Andlauer, T. M. F., Bacanu, S. A., Baekvad-
1050 Hansen, M., Beekman, A. F. T., Bigdeli, T. B., Binder, E. B., Blackwood, D. R. H.,
1051 Bryois, J., Buttenschon, H. N., Bybjerg-Grauholm, J., . . . Major Depressive Disorder
1052 Working Group of the Psychiatric Genomics, C. (2018). Genome-wide association
1053 analyses identify 44 risk variants and refine the genetic architecture of major
1054 depression. *Nat Genet*, *50*(5), 668-681. <https://doi.org/10.1038/s41588-018-0090-3>
- 1055 Xia, C. H., Ma, Z., Ciric, R., Gu, S., Betzel, R. F., Kaczkurkin, A. N., Calkins, M. E., Cook, P.
1056 A., Garcia de la Garza, A., Vandekar, S. N., Cui, Z., Moore, T. M., Roalf, D. R.,
1057 Ruparel, K., Wolf, D. H., Davatzikos, C., Gur, R. C., Gur, R. E., Shinohara, R. T., . . .
1058 Satterthwaite, T. D. (2018). Linked dimensions of psychopathology and connectivity
1059 in functional brain networks. *Nat Commun*, *9*(1), 3003.
1060 <https://doi.org/10.1038/s41467-018-05317-y>
- 1061 Xu, L., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). Modeling tabular
1062 data using conditional gan. *Advances in neural information processing systems*, *32*.

1063

1064

1065 **Supplementary Materials**

1066 **Details on Genotype Data**

1067 The saliva DNA samples of study participants were collected, and 733,293
1068 single nucleotide polymorphisms (SNPs) were genotyped at Rutgers University Cell
1069 and DNA Repository (RUCDR) with Affymetrix NIDA Smoke Screen Array. Using
1070 PLINK 1.90, we excluded SNPs with genotype call rate <95%, sample call rate
1071 <95%, and minor allele frequency (MAF) <1% before imputation. The genotypes
1072 were imputed using the Michigan Imputation Server (Das et al., 2016) using the 1000
1073 Genome phase3 version5 panel (Genomes Project et al., 2015) with Eagle v2.4
1074 phasing (Loh et al., 2016). Then, the imputed variants with INFO score > .3 that did
1075 not meet our quality control criteria (i.e., call rate <95%, MAF <1%, and Hardy–
1076 Weinberg equilibrium p-value <1e-6) were additionally filtered out. To address
1077 potential bias derived from genetically diverse and related family members in the
1078 ABCD study, we employed PC-Air (Conomos et al., 2015) and PC-Relate (Conomos
1079 et al., 2016) to obtain genetically unrelated individuals beyond 4th-degree relatives
1080 (i.e., kinship coefficient >0.022) and to remove outliers beyond 6 SD limits from the
1081 center of ancestrally informative principal component (PC) space. After quality
1082 control procedures, we included a total of 11,301,999 variants in 10,199 multiethnic
1083 participants. From these, we excluded first-, second-, or third-degree related
1084 samples, resulting in 8,620 unrelated multiethnic participants, among whom 6,555
1085 were of European ancestry.

1086

1087 **Details on Polygenic Scores (PGSs)**

1088 Polygenic scores (PGSs) analyzed in this study are the same PGSs data
1089 used in (Joo et al., 2024). PGSs were derived using summary statistics from publicly
1090 available genome-wide association studies (GWAS). Thirty traits were selected
1091 based on relevance to cognitive, psychiatric, and behavioral outcomes. The chosen
1092 GWAS include attention-deficit/hyperactivity disorder (ADHD) (Demontis et al.,
1093 2019), cognitive performance (CP) (Lee et al., 2018), educational attainment (EA)
1094 (Lee et al., 2018), major depressive disorder (MDD) (Wray et al., 2018), insomnia
1095 (Jansen et al., 2019), snoring (Jansen et al., 2019), intelligence quotient (IQ)
1096 (Savage et al., 2018), post-traumatic stress disorder (PTSD) (Nievergelt et al., 2019),
1097 depression (DEP) (Howard et al., 2019; Shen et al., 2020), body mass index (BMI)
1098 (Akiyama et al., 2017; Locke et al., 2015), alcohol dependence (ALCDEP) (Walters
1099 et al., 2018), autism spectrum disorder (ASD) (Grove et al., 2019), automobile
1100 speeding propensity (ASP) (Akiyama et al., 2017), bipolar disorder (BIP) (Stahl et al.,
1101 2019), cannabis during lifetimes (Cannabis) (Pasman et al., 2019), ever smoker
1102 (Karlsson Linner et al., 2019), shared effects on five major psychiatric disorder
1103 (CROSS) (Cross-Disorder Group of the Psychiatric Genomics, 2013), alcoholic
1104 drinks consumption per week (Drinking) (Karlsson Linner et al., 2019), anorexia
1105 nervosa (Watson et al., 2019), neuroticism (Nagel et al., 2018), obsessive-
1106 compulsive disorder (OCD) (International Obsessive Compulsive Disorder
1107 Foundation Genetics & Studies, 2018), first principal components of four risky
1108 behaviors (Risky Behav) (Karlsson Linner et al., 2019), general risk tolerance
1109 (RiskTol) (Karlsson Linner et al., 2019), schizophrenia (SCZ) (Bipolar et al., 2018;
1110 Lam et al., 2019), worrying (Nagel et al., 2018), anxiety (Otowa et al., 2016),
1111 subjective well-being (SWB) (Okbay et al., 2016), general happiness (UK Biobank
1112 GWAS. Neale Lab. <http://www.nealelab.is/ukbiobank/>), and general happiness for

1113 health (happiness-health) (UK Biobank GWAS. Neale Lab.
1114 <http://www.nealelab.is/ukbiobank/>) and meaningful life (happiness-meaning) (UK
1115 Biobank GWAS. Neale Lab. <http://www.nealelab.is/ukbiobank/>).

1116 For 25 of these traits, where only European-based GWAS were available, we
1117 calculated polygenic scores using European-based GWAS summary statistics. For
1118 five traits where multiethnic GWAS results were available (PTSD, DEP, BMI,
1119 ALCDEP, SCZ), we calculated polygenic scores using both European-based and
1120 multiethnic-based GWAS summary statistics. In the European-only analysis, we
1121 used polygenic scores for all 30 traits calculated based on European-based GWAS.

1122 The GWAS summary statistics were used as input for PRS-CS (Ge et al.,
1123 2019), a Bayesian regression method, to estimate the posterior effect sizes of SNPs.
1124 The final scores were calculated using PLINK v1.9. To optimize the scores, we
1125 followed the suggestion of the original PRS-CS paper and chose the optimal global
1126 shrinkage hyperparameter (ϕ) from among four possible values: 1, 1e-2, 1e-4,
1127 and 1e-6. The validation procedure was carried out within 14 PGSs (i.e., DEP, MDD,
1128 ADHD, general happiness, happiness-health, happiness-meaning, SWB, insomnia,
1129 snoring, BMI, PTSD, CP, EA, IQ) that had related measures in the ABCD study. For
1130 each PGS, we performed linear regression of the phenotype variable with each of
1131 the four scores and covariates (sex, age, and the first ten genetic PCs), and then,
1132 based on R^2 and beta coefficient of PGS, selected one of the four PGSs. The
1133 remaining 16 PGSs was automatically validated by PRS-CS-auto (Ge et al., 2019),
1134 which select the optimal value of global shrinkage parameter employing a Bayesian
1135 approach. Finally, to minimize the bias from population stratification, we residualized
1136 the final PGSs with the first ten genetic PCs.

1137 **Supplementary Table 1. Demographic information of the**
 1138 **supplementary analysis participants (multiethnic analysis and**
 1139 **complete data samples).**

		PGS – BNM (n=7,297)		ENV – BNM (n=4,215)		Pheno – BNM (n=6,598)	
		train	test	train	test	train	test
N	total	5,837	1,460	3,372	843	5,278	1,320
Sex	Male	3,087	765	1,790	440	3,031	758
	Female	2,747	695	1,581	403	2,247	562
	Other	3	0	1	0	0	0
Race	White	3,251	780	1,702	451	3,036	765
	Black	798	207	474	103	674	161
	Hispanic	1,176	302	781	183	933	228
	Asian	18	6	86	21	94	19
	Other	594	165	329	85	541	147

1140

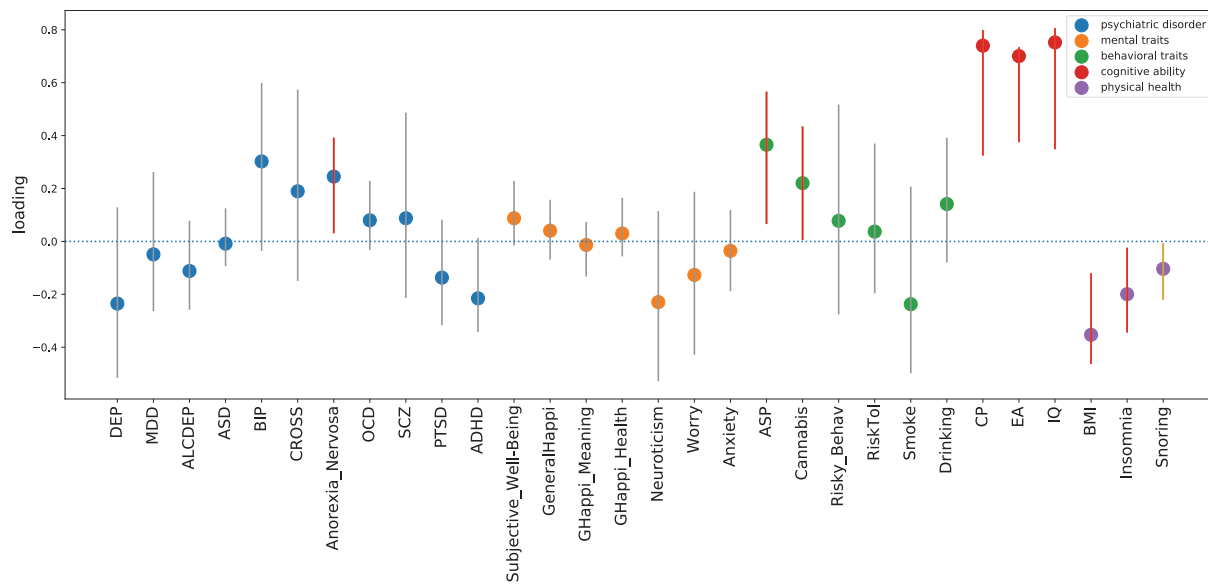
1141

1142

1143

1144

1145 **Supplementary Figure 1. SCCA Loadings of Polygenic Scores in**
1146 **European Samples.**



1147

1148

1149

1150

1151

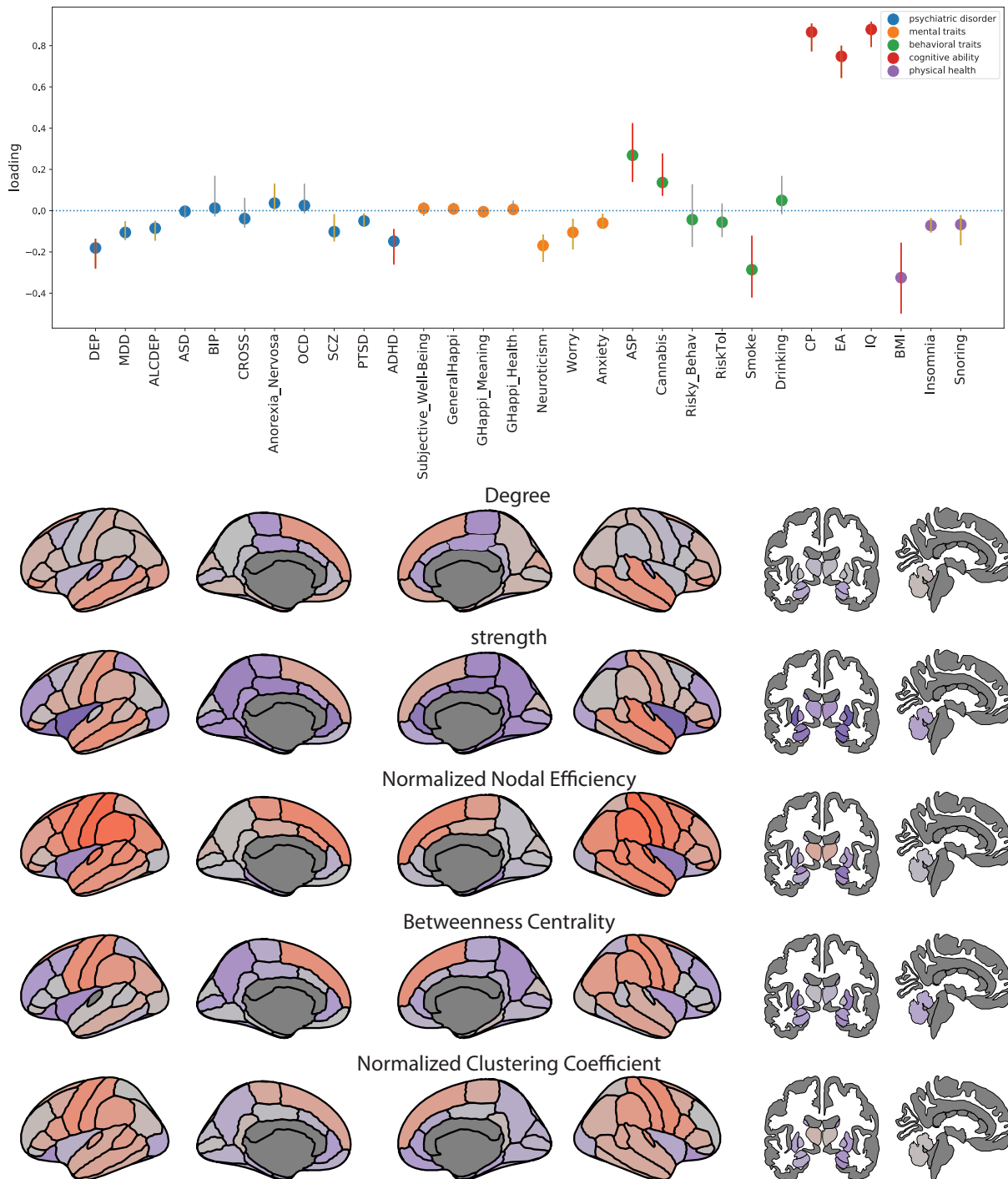
1152

1153

1154

1155 **Supplementary Figure 2. SCCA Loadings of Polygenic Scores and**

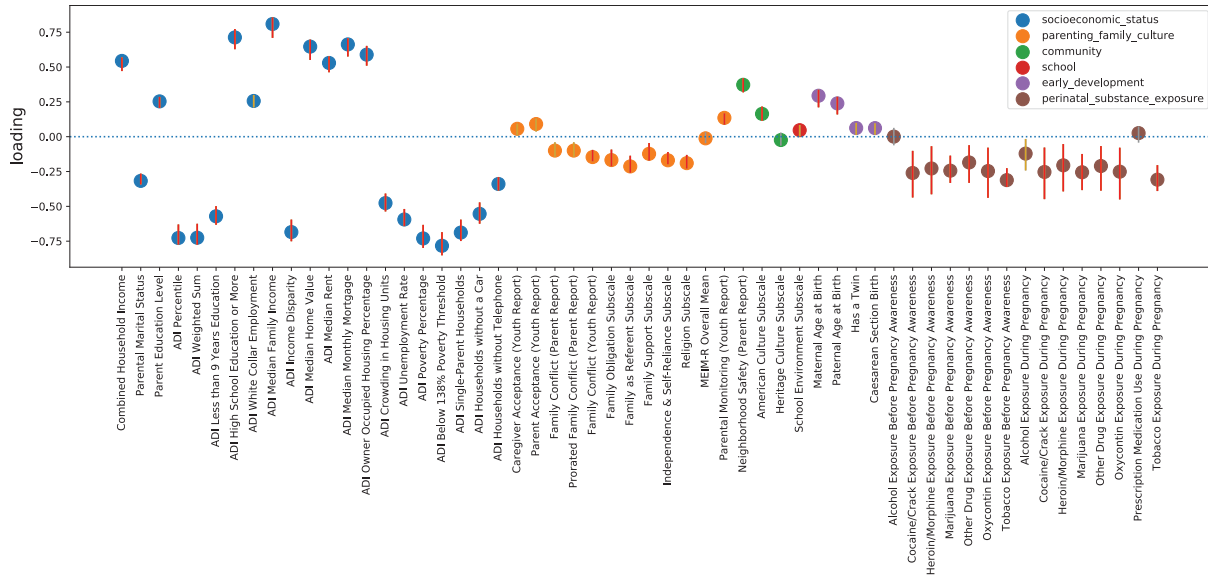
1156 **Brain Network Measures in Multiethnic Samples.**



1157

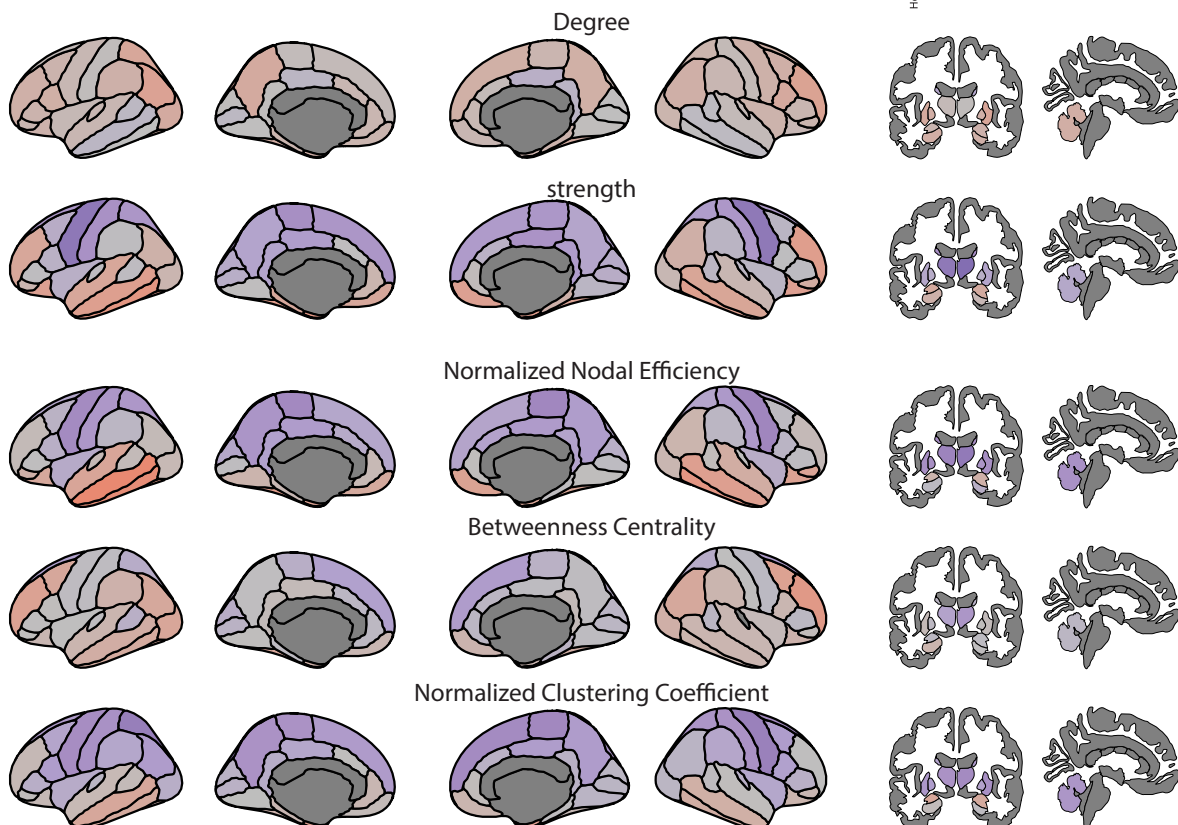
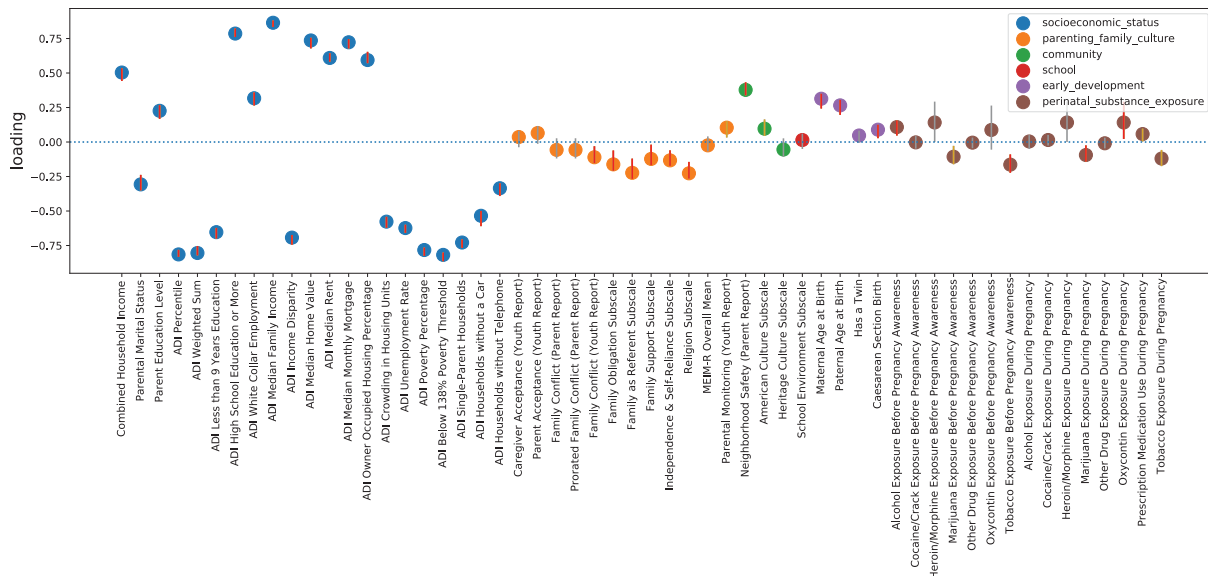
1158

1159 **Supplementary Figure 3. SCCA Loadings of Environmental**
 1160 **Variables in the Main Analysis.**



1161
 1162
 1163
 1164
 1165
 1166
 1167
 1168
 1169

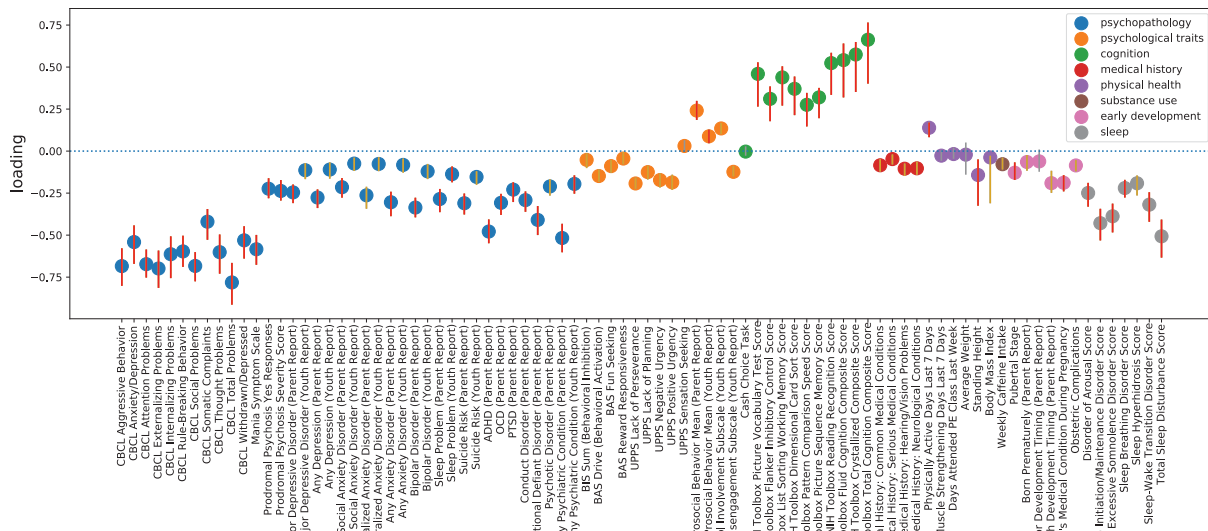
1170 **Supplementary Figure 4. SCCA Loadings of Environmental**
 1171 **Variables and Brain Network Measures in Non-Imputed Samples.**



1172

1173

1174 **Supplementary Figure 5. SCCA Loadings of Phenotype Variables in**
 1175 **the Main Analysis Mode 1.**



1176

1177

1178

1179

1180

1181

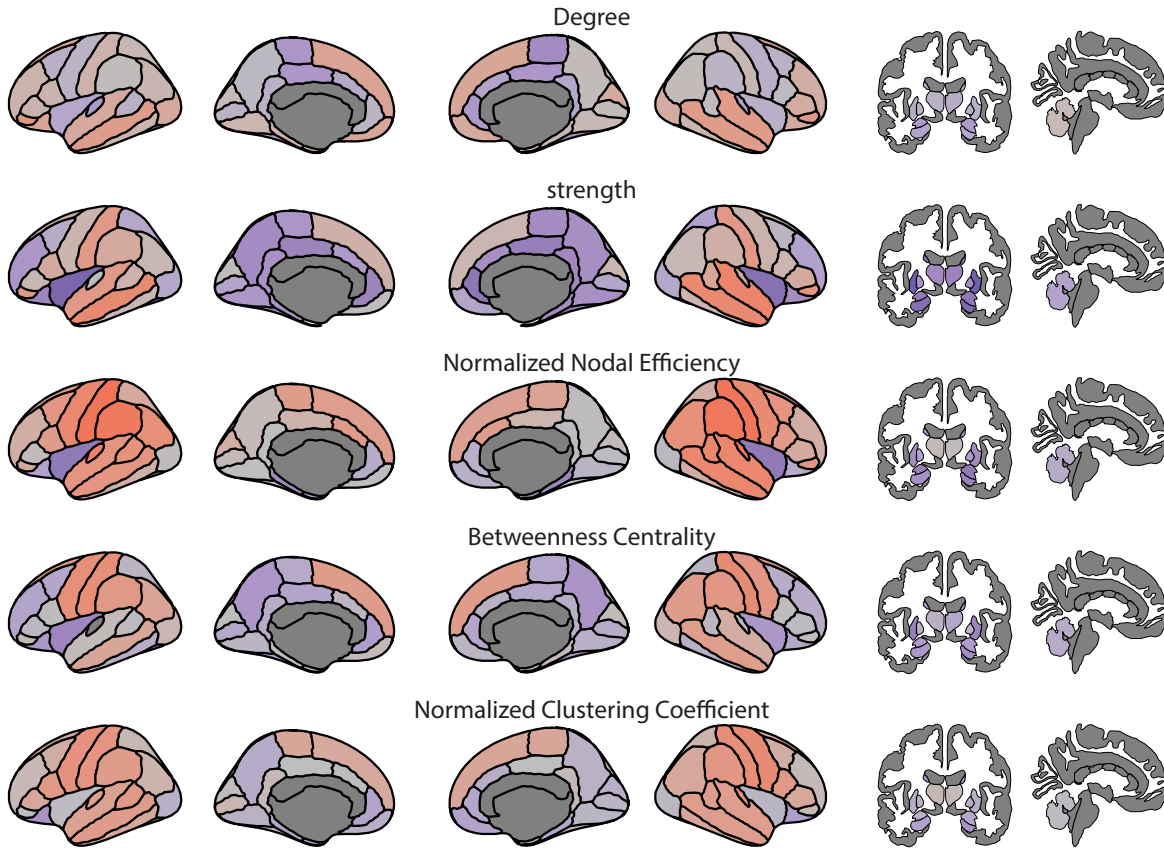
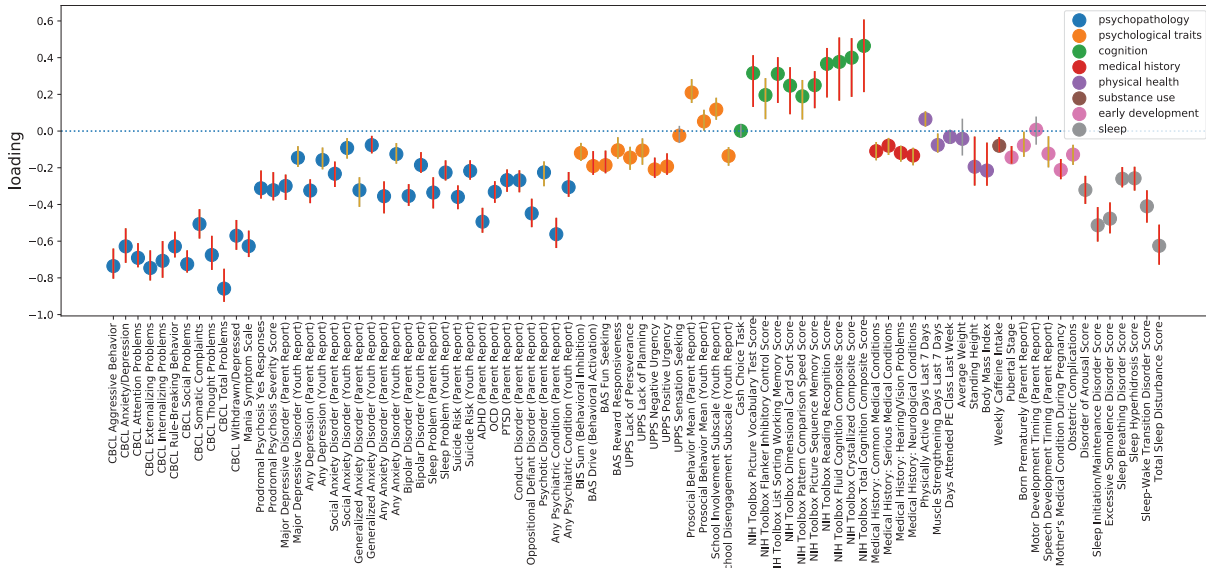
1182

1183

1184

1185 **Supplementary Figure 6. SCCA Loadings of Phenotype Variables**

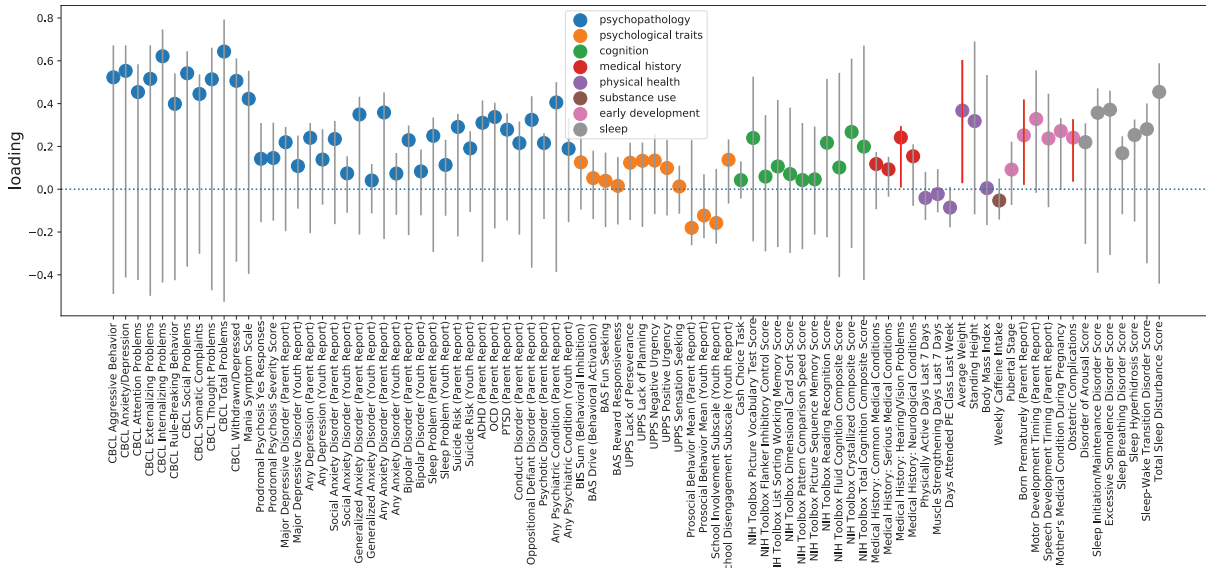
1186 **and Brain Network Measures in Non-Imputed Samples Mode 1.**



1187

1188

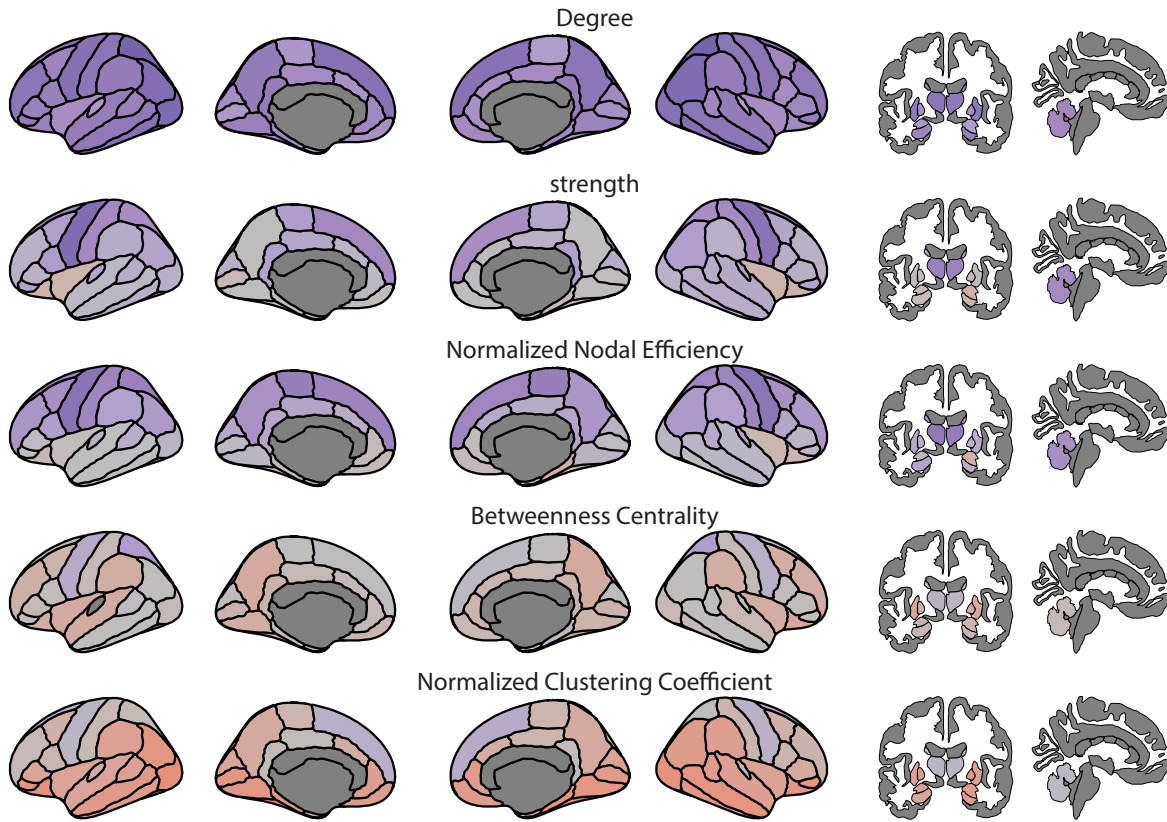
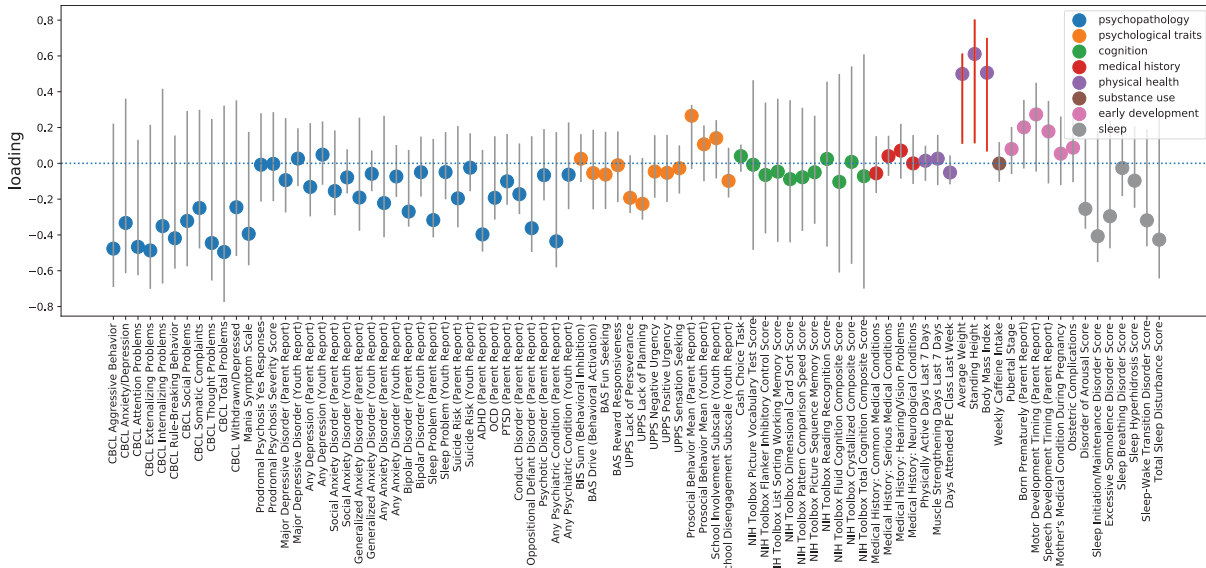
1189 **Supplementary Figure 7. SCCA Loadings of Phenotype Variables in**
1190 **the Main Analysis Mode 2.**



1191
1192
1193
1194
1195
1196
1197
1198
1199

1200 **Supplementary Figure 8. SCCA Loadings of Phenotype Variables**

1201 **and Brain Network Measures in Non-Imputed Samples Mode 2.**



1202