Bridging brain and cognition: A multilayer network analysis of brain structural covariance and general intelligence in a developmental sample of struggling learners

Ivan L. Simpson-Kent[1, *], Eiko I. Fried[2], Danyal Akarca[1, ┼], Silvana Mareva[1, ┼], Edward T. Bullmore[3], the CALM Team and Rogier A. Kievit[1, 4]

[1] MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, Cambridgeshire, CB2 7EF, UK

[2] Department of Clinical Psychology, Leiden University, Leiden, The Netherlands

[3] Department of Psychiatry, University of Cambridge Clinical School, Cambridge, UK

[4] Cognitive Neuroscience Department, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands

***** Correspondence: Ivan.Simpson-Kent@mrc-cbu.cam.ac.uk; 15 Chaucer Road, Cambridge CB2 7EF, UK; Tel.: +44 (0) 1223 769899

┼ Equal contributions

ABSTRACT

Network analytic methods that are ubiquitous in other areas, such as systems neuroscience, have recently been used to test network theories in psychology, including intelligence research. The network or mutualism theory of intelligence proposes that the statistical associations among cognitive abilities (e.g. specific abilities such as vocabulary or memory) stem from causal relations among them throughout development. In this study, we used network models (specifically LASSO) of cognitive abilities and brain structural covariance (grey and white matter) to simultaneously model brain-behavior relationships essential for general intelligence in a large (behavioral, N=805; cortical volume, N=246; fractional anisotropy, N=165), developmental (ages 5-18) cohort of struggling learners (CALM). We found that mostly positive, small partial correlations pervade both our cognitive and neural networks. Moreover, calculating node centrality (absolute strength and bridge strength) and using two separate community detection algorithms (Walktrap and Clique Percolation), we found convergent evidence that subsets of *both* cognitive and neural nodes play an intermediary role *between* brain and behavior. We discuss implications and possible avenues for future studies.

Keywords: General intelligence, Cortical volume, Fractional anisotropy, Brain structural covariance, Cognitive network neuroscience, Multilayer network analysis

INTRODUCTION

General intelligence or *g* (Spearman, 1904) captures cognitive ability across a variety of domains and predicts a wide range of important life outcomes such as educational and occupational achievement (Hegelund et al., 2018), and mortality (Calvin et al., 2011). As individual differences in intelligence found in childhood remain surprisingly stable across the lifespan (Deary et al., 2004), understanding the emergence of cognitive ability is crucial to supporting individuals, especially those considered 'low-performing' (e.g. students struggling to learn in school). Recent work in network analysis has shed new light on both the cognitive abilities that make up general intelligence (Kievit et al., 2019; Van Der Maas et al., 2017), as well as the brain systems purported to support them (Girn et al., 2019; Seidlitz et al., 2018).

In psychological sciences, factor models have traditionally been used to study intelligence (e.g. Carroll, 1993). However, in the last two decades there has been a rise in use of the statistical tools of network science (Barabási, 2016) to show that examining relationships between cognitive abilities (e.g. reasoning and vocabulary) can help us better understand the *development* of general intelligence. For instance, the mutualism model (Van Der Maas et al., 2006) was inspired by an ecosystem model of prey-predator relations, and states that the positive manifold (Spearman, 1904) *emerges* gradually from the positive interactions among different cognitive abilities (e.g. reasoning and vocabulary) over time (see Kievit et al., 2017; 2019). Hence, *g* can arise even from originally weakly correlated cognitive faculties. This highlights the need to both conceptualize traits, abilities, or psychological constructs such as general intelligence as complex dynamical systems, as well as use appropriate statistical models widely used in other disciplines (i.e. network analysis) to estimate relationships among elements of the systems under investigation (Fried, 2020; Fried & Robinaugh, 2020).

Although longitudinal data are preferable to study developmental questions (e.g. Kievit et al., 2013; Rhemtulla et al., 2020), cross-sectional network analysis can provide important

insights. Recent innovations in network psychometrics (Epskamp et al., 2018) have led to a rapid increase in popularity of behavioral network analysis, especially in psychopathology (Borsboom, 2017; Robinaugh et al., 2019). In this framework, psychological constructs are theorized as complex systems and relationships (edges) between nodes (e.g. item responses on a questionnaire) are estimated using weighted partial correlation networks, which enable determination of conditional dependencies among variables after controlling for the associations among every other node in the network (Epskamp et al., 2018).

Recently, this approach has also been used to analyze cross-sectional data on general intelligence. For instance, both Kan et al., 2019 (N=1,800; age range: 16-89 years) as well as Schmank et al., 2019 (N=1,112; age range: 12-90 years) found support for mutualism in the WAIS-IV cognitive battery (Wechsler, 2008), such that a network model showed better fit to the pattern of intelligence scores compared to a latent variable approach (*g*-factor). Lastly, Mareva & Holmes, 2020, in two separate samples, one the same group of struggling learners as studied here (see Holmes et al., 2019) but with less participants (N=350), observed links between cognitive abilities and learning, especially between mathematics skills and more "domain-general" faculties such as backward digit span and matrix reasoning.

In neuroscience, network analysis methods have been widely used to describe the relations among brain regions, ushering in the field of network neuroscience (Bassett & Sporns, 2017; Fornito et al., 2016). Rather than focusing on individual brain regions in isolation, the brain is conceived as a complex system of interconnected networks that facilitate behavioral functions ranging from sensorimotor control to learning. In this light, several influential studies have revealed pervasive properties of brain networks such as small-world topology (Bassett & Bullmore, 2006; 2017), modularity (Meunier et al., 2010; Sporns & Betzel, 2016) and 'rich-club' connector hubs (Heuvel & Sporns, 2011), consistent with an economical trade-off

between minimizing wiring cost and maximizing efficiency (e.g. information transfer) that enable adaptive behavior (Bullmore & Sporns, 2012).

One proposal attempting to explain general intelligence using network neuroscience is The Network Neuroscience Theory of Human Intelligence (Barbey, 2018). Barbey argues that general intelligence arises from the dynamic small-world typology of the brain, which permits transitions between "regular" or "easy-to-reach" network states (needed to access prior knowledge for specific abilities) and "random" or "difficult-to-reach" (required to integrate information for broad abilities) network states (i.e. as in network control theory, see Gu et al., 2015). Together, this constrained flexibility allows the brain to adapt to novel cognitive domains (e.g. in abstract reasoning) while still preserving access to previously learned skills (e.g. from schooling).

Although network approaches have provided unique insights within neuroscience as well as cognitive psychology, few studies have integrated them into a so-called multilayer network paradigm (Bianconi, 2018), which models the relationships among variables simultaneously across scales (e.g. days weeks and years) or domains (e.g. behavior and brain variables). One study has recently pushed this boundary. Hilland et al., 2020 examined the relations between brain structure (cortical thickness and volume) and depression symptoms. They found (via a partial correlation network model) that certain clusters of brain regions (cingulate, fusiform gyrus, hippocampus, and insula) were conditionally dependent with a subset of depression symptoms (crying, irritability and sadness).

This study builds on these findings, by combining a network psychometrics approach to understand individual differences in cognitive ability with structural covariance networks derived from structural brain properties (grey matter cortical volume and white matter fractional anisotropy). The advantages of applying the *networks of networks* approach (Kenett et al., 2015) are threefold and complementary. First, it allows for easier visualizations of

potential causal links between cognition and structural brain properties, which to our knowledge, has only been done in a similar way in one other study involving depression (Hilland et al., 2020). Second, it enables the use of community detection algorithms to tease apart major clusters of cognitive abilities, which could help pinpoint potential intervention targets (e.g. using cognitive training and/or transcranial magnetic stimulation). Lastly, it aids in establishing a coherent framework for theory building, which has been lacking in both the neuroscience (Levenstein et al., 2020) and psychological (Fried, 2020) literature. Ultimately, the hope is that relations between brain-behavior nodes can help identify candidate targets (e.g. nodes that bridge the brain and cognition) for future interventions in developmental samples of struggling learners.

METHODS

Participants

The present sample (behavioral, N=805; cortical volume, N=246; fractional anisotropy, N=165; age range: 5 to 18 years) was obtained from the Centre for Attention, Learning and Memory (CALM) located in Cambridge, UK (Holmes et al., 2019). This developmental cohort consists of children and adolescents recruited by referrals for perceived difficulties in attention, memory, language, reading and/or mathematics problems. A formal diagnosis was neither required nor an exclusion criterion. Exclusion criteria included any known significant and uncorrected problems in vision or hearing, and/or being a non-native English speaker.

Cognitive data were obtained on a one-to-one basis by an examiner in a designated child-friendly testing room. The tasks analyzed in this study comprised a comprehensive array of standardized assessments of cognitive ability including crystallized intelligence (peabody picture vocabulary test, spelling, single word reading, and numerical operations), fluid intelligence (matrix reasoning), and working memory (forward and backward digit recall, Mr.

X, dot matrix, and following instructions). See Table 1 for tasks descriptions, relevant citations, and summary statistics.

Cognitive Domain	Task Descriptions	Mean (sd) [range]	Missing Data	Reference
Crystallized Ability (gc)	**Numerical Operations (NO):** Participants answered written mathematical problems that increased in difficulty.	16.1 (8.4) [0, 64]	9.94%	Wechsler, 2005
	Single Word Reading (Read): Participants read aloud first a list of letters and then words that gradually increased in complexity. Correct responses required correctness and fluency.	83.2 (24.8) [1, 130]	2.48%	
	Spelling (Spell): Participants spelled words with increasing difficulty one at a time that were spoken by an examiner.	22 (9.2) [0, 49]	3.35%	
	Peabody Picture Vocabulary Test (Pea): Participants were asked to choose the picture (out of four multiple-choice options) showing the meaning of a word spoken by an examiner.	136.8 (31.6) [8, 215]	1.12%	Dunn and Dunn, 2007
Fluid Ability (gf)	**Matrix Reasoning (MR):** Participants saw sequences of partial matrices and selected the response option that best completed each matrix.	11.2 (5.6) [0, 28]	0.12%	Wechsler, 2011
Working Memory (WM)	**Digit Recall (DR):** Participants recalled sequences of single digit numbers given in audio format.	24.6 (5.4) [7, 47]	0.5%	Alloway, 2007
	Backward Digit Recall (BDR): Same as regular digit recall but in reversed order.	9.7 (4.4) [0, 25]	3.11%	
	Dot Matrix (Dot): Participants were shown the location of a red dot in a sequence of 4x4 matrices and had to recollect the location and order of these sequences.	18 (5.7) [2, 43]	0.75%	

	Mr. X (MrX): Participants remembered spatial sequences of locations of a ball held by a cartoon man rotated in one of seven positions.	9.3 (5.1) [0, 32]	1.24%	
	Following Instructions (FI): Participants carried out various sequences of actions (touch and/or pick up) based on props (a box, an eraser, a folder, a pencil or a ruler) presented in front of them. By having participants undertake actions sequentially (do X "then" do Y), increasingly longer sequences were made which increased the difficulty. Scores denote total number of correct responses.	11.2 (4) [1, 33]	6.83%	Gathercole et al., 2008

Table 1. List, descriptions, and summary statistics (mean, standard deviation, range, and percentage of missing data) of cognitive assessments used in the study from the CALM sample. Note, tasks descriptions (except following instructions) are taken directly or paraphrased from Simpson-Kent et al., 2020.

Participants were allotted regular breaks throughout each session. When necessary, testing was split into two separate sessions for participants who did not complete the assessments in a single sitting. A subset of participants also underwent MRI scanning (see below for details).

Structural neuroimaging: cortical volume (CV) and fractional anisotropy (FA)

CALM neuroimaging data were obtained at the MRC Cognition and Brain Sciences Unit, Cambridge, UK. Scans were acquired on the Siemens 3 T Tim Trio system (Siemens Healthcare, Erlangen, Germany) via 32-channel quadrature head coil. T1-weighted volume scans were acquired using a whole brain coverage 3D magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence, with 1 mm isotropic image resolution. The following parameters were used: Repetition time (TR) = 2250 ms; Echo time (TE) = 3.02 ms; Inversion

time (TI) = 900 ms; Flip angle = 9 degrees; Voxel dimensions = 1 mm isotropic; GRAPPA acceleration factor = 2. Diffusion-Weighted Images (DWI) were acquired using a Diffusion Tensor Imaging (DTI) sequence with 64 diffusion gradient directions, with a b-value of 1000 s/mm^2, plus one image acquired with a b-value of 0. Relevant parameters include: TR = 8500 ms, TE = 90 ms, voxel dimensions = 2 mm isotropic.

We undertook several procedures to ensure adequate MRI data quality and minimize potential biases due to subject movement. For all participants in CALM, children were trained to lie still inside a realistic mock scanner prior to their scan. All T1-weighted images and FA maps were examined by an expert to remove low quality scans. Moreover, only data with a maximum between-volume displacement below 3 mm were included in the analyses.

As our grey matter metric, we use region-based cortical volume (CV in mm^3, N=246, averaged across contralateral homologues), based on the Desikan-Killiany atlas (Desikan et al., 2006) and defined as the distance between the outer edge of cortical grey matter and subcortical white matter (Bruce Fischl & Dale, 2000). Tissue classification and anatomical labelling was performed on the basis of the T1-weighted scan using FreeSurfer v5.3.0 software which is documented and freely available for download online (http://surfer.nmr.mgh.harvard.edu/). The technical details of these procedures are described in prior publications (Dale et al., 1999; Fischl et al., 1999; 2002). FreeSurfer morphology output statistics were computed for each ROI, and also included cortical thickness and surface area (see Supplementary Material). Based on a recent meta-analyses on functional and structural correlates of intelligence (Basten et al., 2015), we included a subset of 10 cortical volume regions in this study: caudal anterior cingulate (CAC), caudal middle frontal gyrus (CMF), frontal pole (FP), medial orbital frontal cortex (MOF), rostral anterior cingulate gyrus (RAC), rostral middle frontal gyrus (RMF), superior frontal gyrus (SFG), superior temporal gyrus (STG), supramarginal gyrus (SMG), and transverse temporal gyrus (TTG).

From a subset our neuroimaging data (see Simpson-Kent et al., 2020), we also calculated fractional anisotropy (FA, N=165), a proxy for white matter integrity (Wandell, 2016). We included 10 regions using the Johns Hopkins University DTI-based white matter tractography atlas (see Hua et al., 2008): anterior thalamic radiations (ATR), corticospinal tract (CST), cingulate gyrus (CING), cingulum [hippocampus] (CINGh), forceps major (FMaj), forceps minor (FMin), inferior fronto-occipital fasciculus (IFOF), inferior longitudinal fasciculus (ILF), superior longitudinal fasciculus (SLF), and uncinate fasciculus (UNC).

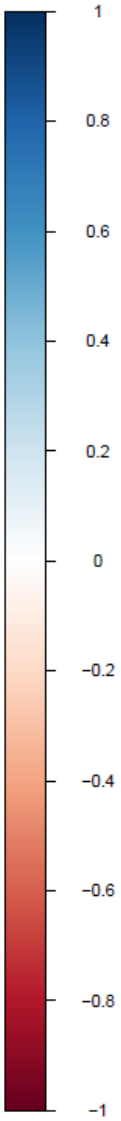
All steps to compute regional CV estimation and FA maps were implemented using NiPyPe v0.13.0 (see https://nipype.readthedocs.io/en/latest/). To create a brain mask based on the b0-weighted image (FSL BET; Smith, 2002) and correct for movement and eddy current-induced distortions (eddy; Graham et al., 2016), diffusion-weighted images were pre-processed. The diffusion tensor model was then fitted and fractional anisotropy (FA) maps were calculated using *dtifit*. Images with a between-image displacement >3 mm were then excluded from subsequent analysis steps. This was completed using FSL v5.0.9. To extract FA values for major white matter tracts. FA images were registered to the FMRIB58 FA template in MNI space using a sequence of rigid, affine, and symmetric diffeomorphic image registration (SyN). This was implemented in ANTS v1.9 (Avants et al., 2008). For all participants, visual inspection indicated good image registration. Binary masks from a probabilistic white matter atlas (thresholded at >50 % probability) in the same space were applied to extract FA values.

We used these region based measures to study brain structural covariance (Alexander-Bloch et al., 2013), which have been used in cross-sectional and longitudinal designs of cognitive ability in childhood and adolescence (e.g. Solé-Casals et al., 2019; see Kievit & Simpson-Kent, 2019 for a review of longitudinal studies). Emerging theoretical proposals emphasize the role of networks of brain areas in producing intelligent behavior (e.g. Parieto-Frontal Integration Theory (P-FIT), Jung & Haier, 2007 and The Network Neuroscience

Theory of Human Intelligence, Barbey, 2018) rather than individual regions-of-interest (ROIs) in isolation (e.g. primarily the prefrontal cortex). As stated above, we selected 10 grey matter and 10 white matter ROIs based upon combined evidence from a recent meta-analysis (Basten et al., 2015) on associations between functional and structural ROIs and cognitive ability that further extended the P-FIT theory, but also more recent work done in two large cohorts, one in longitudinal analysis of the UK Biobank sample (grey matter, Kievit et al., 2018) and another in a the same (cross-sectional) developmental sample (N=551) studied here (white matter, Simpson-Kent et al., 2020).

For a correlation plot of cognitive tasks and neuroimaging measures, see Figure 1. To view age trends of cognitive tasks and structural neuroimaging measures, see Figure 2. Lastly, see Figure 3 for illustrations of ROIs analyzed in this study.

MR																			
0.69	Pea																		
0.57	0.68	Spell																	
0.59	0.72	0.89	Read																
0.69	0.7	0.73	0.72	NO															
0.45	0.57	0.53	0.56	0.5	DR														
0.57	0.62	0.66	0.66	0.69	0.57	BDR													
0.56	0.58	0.5	0.5	0.62	0.4	0.56	MrX												
0.58	0.59	0.55	0.55	0.63	0.46	0.61	0.57	Dot											
0.49	0.51	0.47	0.5	0.51	0.4	0.47	0.44	0.46	Ins										
0.08	0.1	0.1	0.13	0.13	0.21	0.22	0.15	0.16	0.03	CAC									
0.15	0.11	0.09	0.13	0.02	0.15	0.19	0.2	0.22	0.15	0.43	CMF								
−0.04	−0.05	−0.08	0.02	−0.17	0	0.06	0.04	−0.02	0.01	0.4	0.57	MOF							
0.1	0.14	0.2	0.25	0.16	0.2	0.26	0.21	0.22	0.16	0.66	0.51	0.5	RAC						
0.08	0.08	−0.02	0.07	−0.07	0.14	0.11	0.12	0.15	0.1	0.5	0.63	0.67	0.55	RMF					
0.12	0.08	0.08	0.13	0.02	0.1	0.19	0.11	0.18	0.12	0.49	0.66	0.65	0.6	0.84	SFG				
0.14	0.11	0.13	0.19	0.12	0.24	0.23	0.17	0.19	0.15	0.59	0.56	0.57	0.63	0.7	0.74	STG			
0.07	0.12	0.07	0.11	0.08	0.14	0.16	0.16	0.16	0.17	0.45	0.48	0.55	0.54	0.58	0.65	0.69	SMG		
−0.11	−0.12	−0.21	−0.1	−0.22	−0.06	−0.15	−0.05	−0.07	−0.06	0.1	0.15	0.45	0.1	0.44	0.33	0.23	0.14	FP	
0.02	0.01	0.03	0.04	0.02	0.11	0.13	0.03	0.08	0.04	0.41	0.37	0.44	0.35	0.43	0.5	0.68	0.53	0.16	TTG



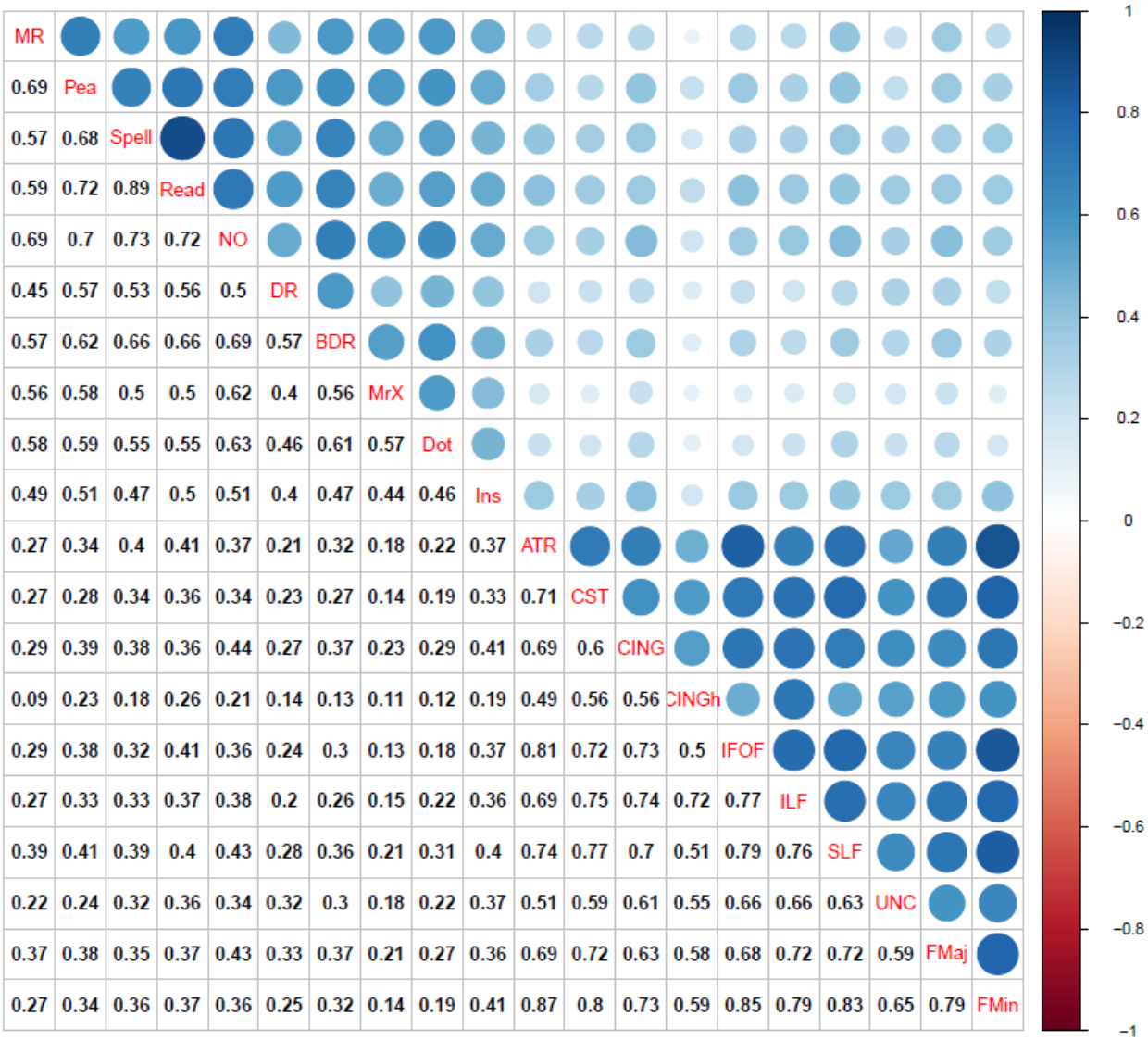
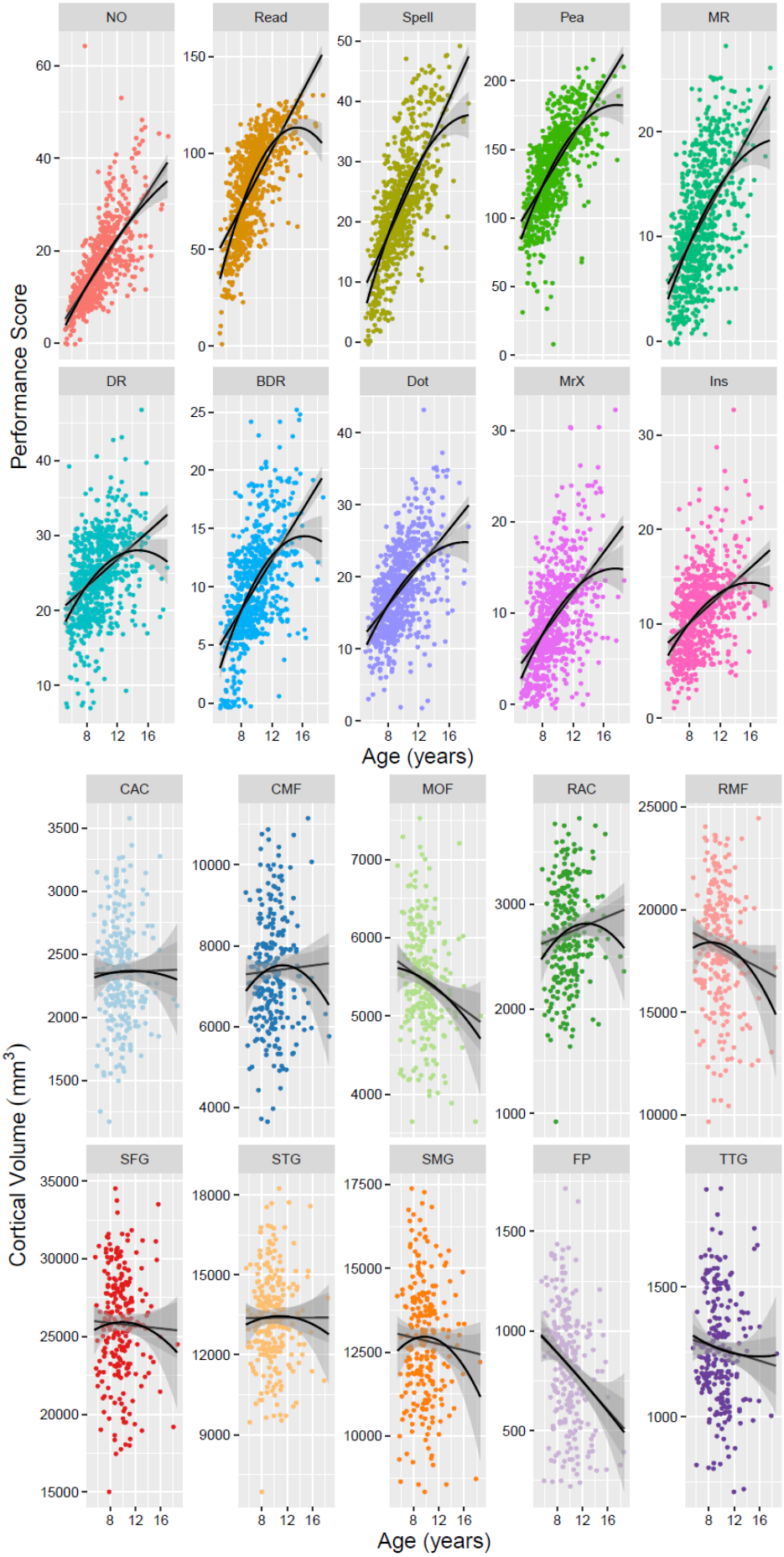


Figure 1. Top: Correlation plot for cognitive raw scores and bilateral cortical volume ROIs. Bottom: Correlation plot for cognitive raw scores and bilateral fractional anisotropy ROIs. Cognitive abbreviations are in Table 1 (above) while neural abbreviations are in section ***Structural neuroimaging: cortical volume (CV) & fractional anisotropy (FA)*** above. All coefficients shown are Pearson correlations. Blue represents positive correlations while red signifies negative correlations among variables. Size of circles indicates the magnitude of the association (e.g. larger circle=higher correlation). Correlations calculated using pairwise complete observations.


NO
Read
Spell
Pea
MR
DR
BDR
Dot
MrX
Ins
Performance Score
Age (years)
CAC
CMF
MOF
RAC
RMF
SFG
STG
SMG
FP
TTG
Cortical Volume (mm3)
Age (years)

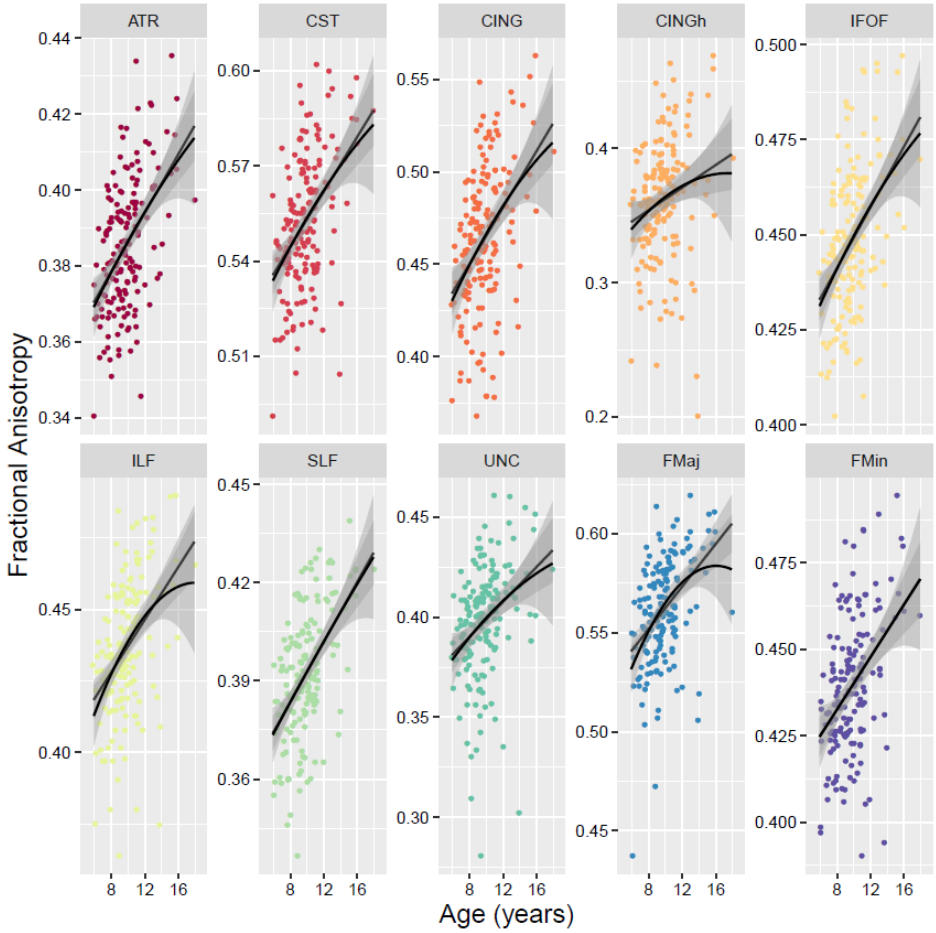


Figure 2. Cross-sectional scatterplots for cognitive raw scores (top), bilateral cortical volume (middle) and bilateral fractional anisotropy (bottom). Solid lines represent linear and polynomial fit while shades indicate 95% confidence intervals.

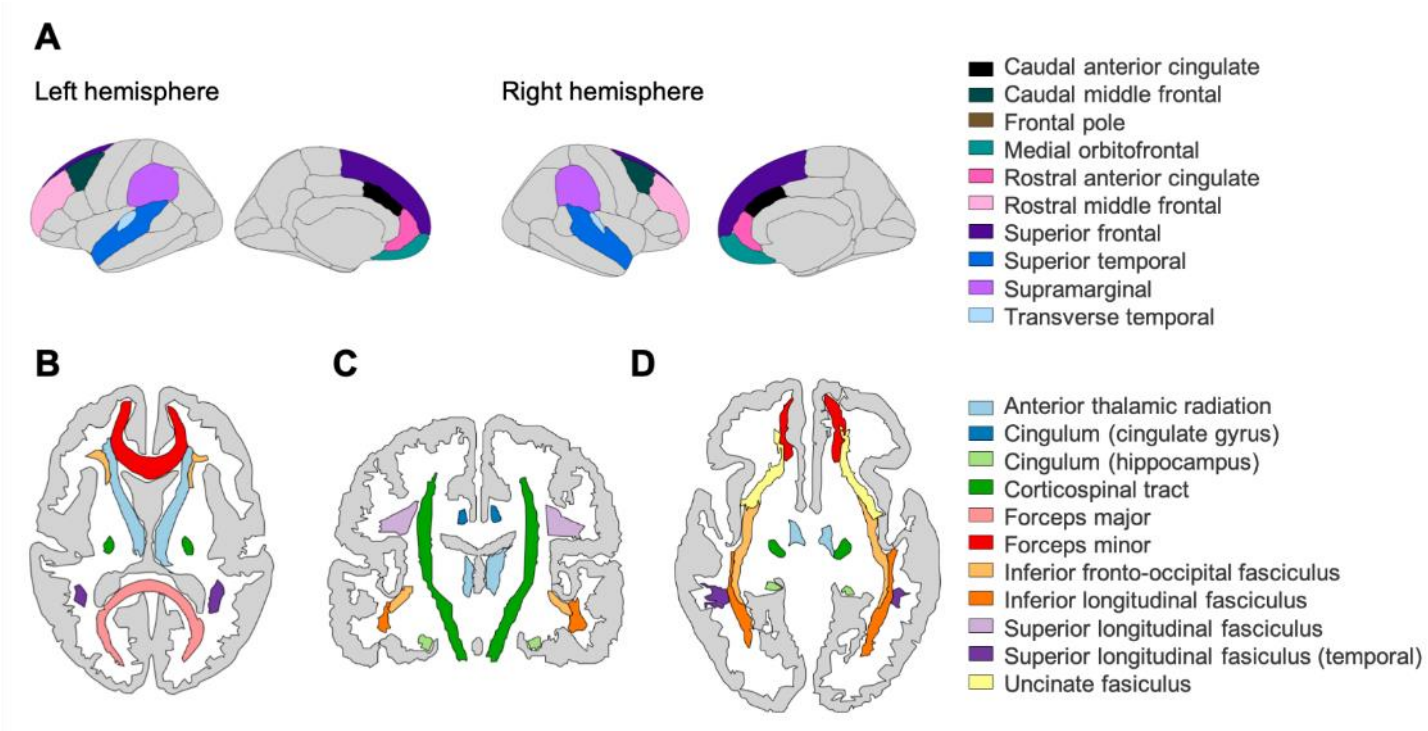


Figure 3. Top: A) Grey matter ROIs based on the DK atlas (cortical volume, N=246) in the left and right hemisphere. Bottom: White matter ROIs based on the John's Hopkin's University atlas (fractional anisotropy, N=165) in B) Transverse plane (superior), C) Coronal plane and D) Transverse plane (inferior).

Network estimation methods

All statistical analyses and plots were completed using R (R Core Team, 2020) version 3.6.3 ("Holding the Windsock"). Network estimation was performed using the packages bootnet (version 1.4.2, Epskamp & Fried, 2020), CliquePercolation (version 0.2.0, Lange & Zickfeld, 2019), igraph (version 1.2.5, Amestoy et al., 2020), qgraph (version 1.6.5, Epskamp et al., 2020), and networktools (version 1.2.3, Jones, 2020). We used these tools to estimate weighted partial correlation networks, which allowed determination of conditional dependencies among our cognitive and neural variables. For example, in a multilayer network, any partial correlation between node A (e.g. matrix reasoning) and node B (e.g. the caudal anterior cingulate) is one that remains *after* controlling for the associations among A and B with every other node in the network (e.g. other cognitive abilities and cortical volume ROIs). To estimate these networks, we applied Gaussian Graphical Models (Pearson correlations)

using regularization (graphical lasso, see Friedman et al., 2008) with a threshold tuning parameter of 0.5 and pairwise deletion to account for missingness. These methods have been widely used to generate sparser networks by penalizing for more complex models—thus, decreasing the risk of potentially spurious (e.g. false positive) connections and enabling simpler visualization and interpretation of conditional dependencies between nodes (Epskamp & Fried, 2018). We hypothesized that our results would show positive partial correlations (in line with mutualism theory) both within cognitive (as observed in Mareva et al., 2019 and Schmank et al., 2019) and within neural measures (single layer networks) as well as between brain-behavior variables in the multilayer networks.

Network centrality

To assess the 'importance' of cognitive and neural nodes relative to others within our networks, we estimated node strength, a weighted degree centrality measure calculated by summing the absolute partial correlation coefficients (edge weights) between a node and all other nodes it connects to within the network. Moreover, our brain structural covariance networks involve ROIs that are not necessarily anatomically connected, preventing certain inferences such as information flow. For all networks, nodes were classified as central if the magnitude of their strength z-score was positive and equal to or greater than one standard deviation above the mean.

For our multilayer networks, instead of traditional strength we calculated bridge strength, a novel weighted degree centrality measure originality developed to study comorbidity between mental disorders (see Jones et al., 2019 for overview) that sums the absolute value of every edge that connects one node (e.g. matrix reasoning) in one pre-assigned community (e.g. cognition) to another node (e.g. caudal anterior cingulate) in another pre-assigned community (e.g. brain). Moreover, this metric has been shown to accurately and

reliably detect true bridge nodes in both directed and undirected networks (Jones et al., 2019). We chose to only measure bridge centrality (strength) for our multilayer networks as excluding this analysis in single-layer networks emphasizes our motivation to hopefully identify potential nodes (both cognitive and neural) that might one day guide intervention studies. We pre-assigned communities for bridge strength calculation based on results from the Walktrap algorithm (see ***Community detection*** below). We did not interpret negative centrality values for our single or multilayer networks.

Next, we quantified the reliability of our centrality estimates for all single-layer (absolute strength of cognitive and brain structural covariance nodes) and multilayer networks (bridge strength). We estimated the correlation stability (CS) coefficient, calculated as the maximum proportion (out of N=2,000 bootstraps) of the sample that can be dropped out and, with 95% probability, still retain a correlation of 0.7, with a CS value of 0.5 considered to be stable (Epskamp et al., 2018). Lastly, we also determined the stability of the edge-weight coefficients but present these results in the Supplementary Material.

Community detection

Finally, to determine in a data-driven manner whether clustering, or grouping, of nodes occurred in our networks, we applied community detection methods separately in our cognitive, neural (grey and white matter), and multilayer networks. We used two commonly used algorithms to increase the robustness of our results: 1) Walktrap (Pons & Latapy, 2005) for all networks and 2) Clique Percolation for weighted networks (Farkas et al., 2007), but *only* for our multilayer networks since they have more concrete pre-defined communities (behavior vs. brain), and, therefore, less probable to overlap in community membership (being 'bridges' between brain and cognition). The presence of a clustering of nodes into communities might further indicate the existence of intermediate endophenotypes (Fornito & Bullmore, 2012;

Kievit et al., 2016) not only among single-layer networks (cognitive or neural) but also (and more importantly for this study) *between* them (e.g. if clusters form consisting of both cognitive and neural nodes that are topologically between distinct communities in our partial correlation networks).

The Walktrap algorithm assesses the similarity between nodes by taking recursive random walks between node pairs and classifies communities according to how densely connected these parts are within the network (wherever the random walks become 'trapped'). Walktrap is widely used in the network psychometrics literature and, in a Monte Carlo simulation study, was shown to outperform other algorithms (e.g. InfoMap) for sparse count networks (e.g. those used in diffusion tensor imaging), although it must be noted that this result was for networks made up of 500 nodes or higher (Gates et al., 2016). We also calculated the maximum modularity index value (Q), which estimates the robustness of the partition (Newman, 2006). We interpreted values of 0.5 or above as evidence for reliable grouping. Finally, Clique Percolation is a method for approximating overlapping community structure in networks (Palla et al., 2005). For example, in the context of this study, a cognitive ability (e.g. matrix reasoning) or brain region (e.g. parts of frontal cortex) might be involved in myriad cognitive (e.g. fluid intelligence vs. working memory, Chuderski, 2015) or neural (e.g. multiple demand system, Assem et al., 2020, vs. P-FIT theory) processes that underlie intelligence. Therefore, Clique Percolation contrasts with the Walktrap algorithm by allowing nodes to be assigned membership to more than one group simultaneously.

RESULTS

Single-layer network models (cognitive, cortical volume and fractional anisotropy)

The regularized partial correlation (PC) network for the CALM cognitive data are shown in Figure 4 (top left). This network shows that all partial correlations are positive, and

most have small magnitude (mean PC=0.08, median PC=0.07, PC range=0—0.63). One edge (between reading and spelling) was an outlier (PC=0.63, all others are between 0 and 0.27), likely due to close content overlap (verbal ability). As this may bias centrality estimates and network interpretation, we implemented several robustness checks described below. Regarding centrality, only one node emerged as strong (positive z-score at or greater than one standard deviation above the mean): reading (Figure 4 top right). Overall, centrality estimates were stable, indicated by a high correlation stability (CS)-coefficient of 0.75 (age included in estimation), revealing that at least 75% of the sample could be dropped while maintaining a correlation of 0.7 with the original sample at 95% probability. As noted above, the highest correlation in the network between spelling and reading (0.63) might biased centrality estimates as both nodes relate to elementary language abilities. Therefore, we re-estimated the network with three modifications: 1) Estimation without the reading node, 2) Estimation without the spelling node, and 3) Estimation after averaging reading and spelling scores into a single node. Re-estimation revealed that the original nodes with positive z-scores were consistent, regardless of modification method. Together, these findings suggest that the strongest nodes in our network are stable.

Finally, we investigated whether the cognitive nodes clustered into distinct communities. The Walktrap algorithm found three cognitive communities (Q=0.23), color coded in Figure 4 (top left): Cluster 1 consisted of matrix reasoning, peabody picture vocabulary test, numerical operations, Mr. X, dot matrix and following instructions; Cluster 2 was made up of reading and spelling; Cluster 3 included forward and backward digit recall.

Next, we estimated the partial correlation network among 10 grey matter regions as shown in Figure 3 (top) above. All edges weights (mean PC=0.09, median PC=0, PC range= -0.15—0.52) of the cortical volume network (Figure 4, middle left) were positive apart from one negative path (caudal middle frontal gyrus and frontal pole PC= -0.15). Note, the negative

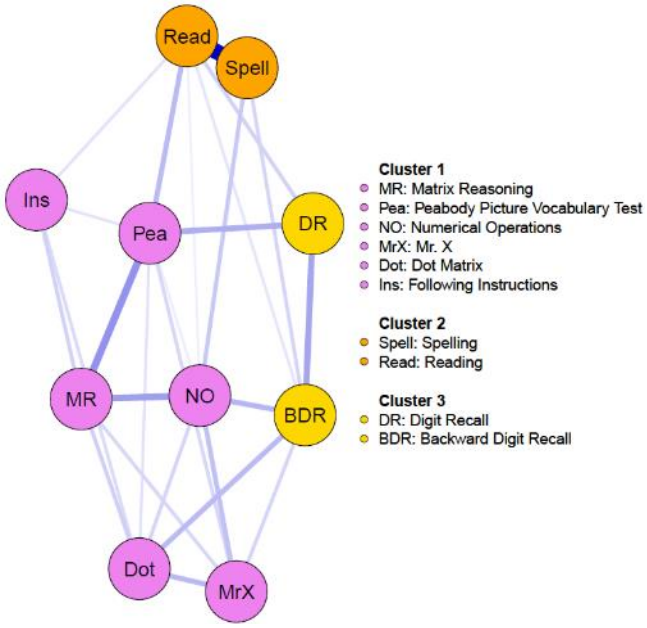
path between the caudal middle frontal gyrus and frontal pole might be due to the frontal pole correlating surprisingly weakly with other grey matter nodes and displaying a steeper decrease pattern across age (Figures 1 and 2).

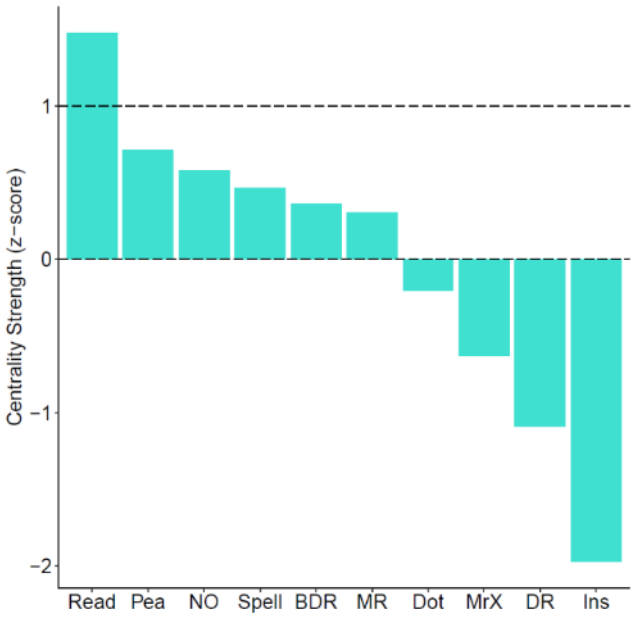
Three ROIs emerged as central (in descending order of centrality strength): superior temporal gyrus, rostral middle frontal gyrus and superior frontal gyrus (Figure 4, middle right). Similar to the cognitive network, cortical volume centrality was stable (CS-coefficient=0.52), indicating that about 52% of the sample could be subtracted to maintain a correlation of centrality estimates above 0.7 compared to the full sample. This finding is despite the lower sample size compared to the behavioral data (N=805 for behavior vs. N=246 for cortical volume).

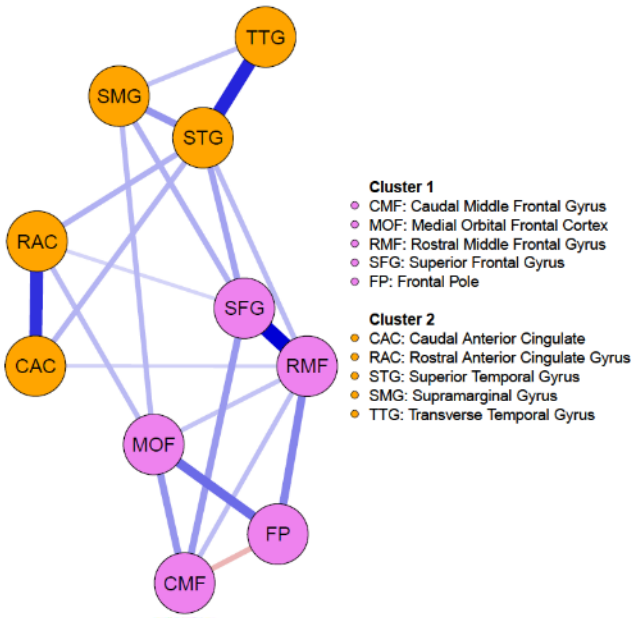
Lastly, to identify possible communities, we again applied the Walktrap algorithm, producing two clusters (Q=0.26): Cluster 1 comprised the caudal middle frontal gyrus, medial orbital frontal cortex, rostral middle frontal gyrus, superior frontal gyrus and frontal pole; Cluster 2 included the caudal anterior cingulate, rostral anterior cingulate gyrus, superior temporal gyrus, supramarginal gyrus and transverse temporal gyrus.

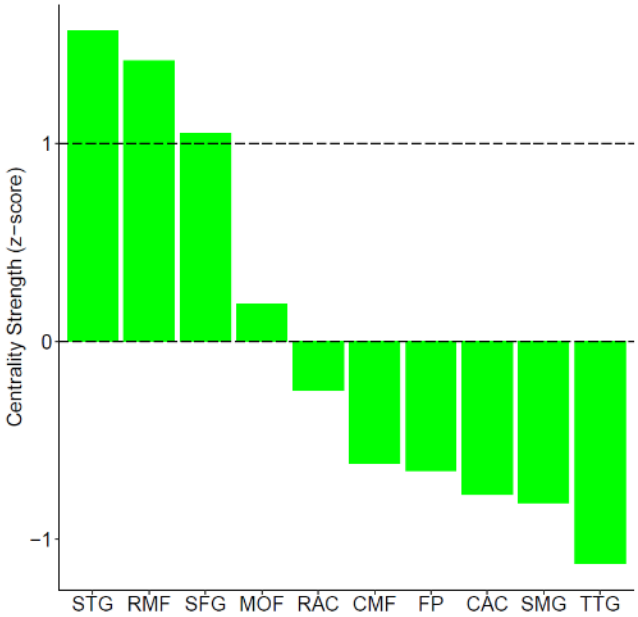
Finally, similar to the cognitive and the grey matter covariance network, the fractional anisotropy network (Figure 4, bottom left) has all positive partial correlations with all edge weights varying between small and moderate values: mean PC=0.08, median PC=0, and PC range=0—0.44. Two white matter ROIs displayed centrality (Figure 4, bottom right). These included (in descending order) the forceps minor and inferior longitudinal fasciculus. Finally, fractional anisotropy centrality was not stable (CS-coefficient=0.36) indicating that only about 36% of the sample could be removed while maintaining a 0.7 association with 95% probability. This is possibly due to the much lower sample size (N=165) compared to the cognitive (N=805) and grey matter (N=246) networks.

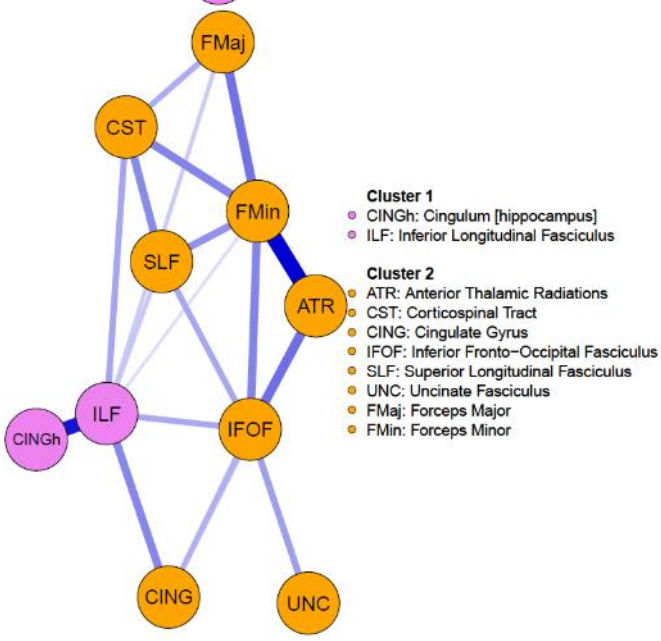
Lastly, using community detection analysis, the Walktrap algorithm generated five ROI groups (Q=0.09): Cluster 1: cingulate gyrus and inferior longitudinal fasciculus; Cluster 2: corticospinal tract, superior longitudinal fasciculus, forceps major and forceps minor; Cluster 3: anterior thalamic radiations and inferior fronto-occipital fasciculus; Cluster 4: hippocampal cingulum; and Cluster 5: uncinate fasciculus. However, closer inspection of the data revealed a single individual with outsized influence in their uncinate fasciculus and forceps major. Removal of this data led to the optimal clustering being only 2 clusters, improving the maximum modularity index (0.16) and did not affect any summary statistics except for range (mean PC=0.08, median PC=0, and PC range= -0.14—0.47): hippocampal cingulum and inferior longitudinal fasciculus (Cluster 1), and cingulate gyrus, corticospinal tract, superior longitudinal fasciculus, forceps major, forceps minor, anterior thalamic radiations, inferior fronto-occipital fasciculus and uncinate fasciculus (Cluster 2). Going forward with our multilayer networks, we include this outlier due to the extensive screening protocol for obtaining the neuroimaging data (see ***Structural neuroimaging: cortical volume (CV) and fractional anisotropy (FA)*** above), but discuss the impact of outliers for all networks (cognitive, neural and multilayer) in the Discussion.


Read
Spell
Ins
Pea
DR
MR
NO
BDR
Dot
MrX
Cluster 1
MR: Matrix Reasoning
Pea: Peabody Picture Vocabulary Test
NO: Numerical Operations
MrX: Mr. X
Dot: Dot Matrix
Ins: Following Instructions
Cluster 2
Spell: Spelling
Read: Reading
Cluster 3
DR: Digit Recall
BDR: Backward Digit Recall


Centrality Strength (z−score)
1
0
−1
−2
Read
Pea
NO
Spell
BDR
MR
Dot
MrX
DR
Ins


TTG
SMG
STG
RAC
SFG
CAC
RMF
MOF
FP
CMF
Cluster 1
CMF: Caudal Middle Frontal Gyrus
MOF: Medial Orbital Frontal Cortex
RMF: Rostral Middle Frontal Gyrus
SFG: Superior Frontal Gyrus
FP: Frontal Pole
Cluster 2
CAC: Caudal Anterior Cingulate
RAC: Rostral Anterior Cingulate Gyrus
STG: Superior Temporal Gyrus
SMG: Supramarginal Gyrus
TTG: Transverse Temporal Gyrus


Centrality Strength (z−score)
1
0
−1
STG
RMF
SFG
MOF
RAC
CMF
FP
CAC
SMG
TTG


FMaj
CST
FMin
SLF
ATR
ILF
CINGh
IFOF
CING
UNC
Cluster 1
CINGh: Cingulum [hippocampus]
ILF: Inferior Longitudinal Fasciculus
Cluster 2
ATR: Anterior Thalamic Radiations
CST: Corticospinal Tract
CING: Cingulate Gyrus
IFOF: Inferior Fronto−Occipital Fasciculus
SLF: Superior Longitudinal Fasciculus
UNC: Uncinate Fasciculus
FMaj: Forceps Major
FMin: Forceps Minor

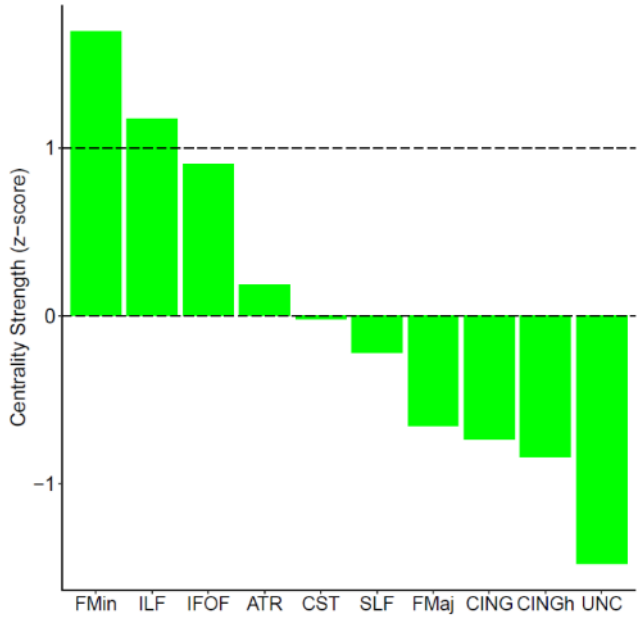

Centrality Strength (z−score)
1
0
−1
FMin
ILF
IFOF
ATR
CST
SLF
FMaj
CING
CINGh
UNC

Figure 4. Single-layer partial correlation networks. Top: Network visualization (spring layout, left side) of CALM cognitive data (N=805). Nodes are grouped according to Walktrap algorithm results (see above). Centrality estimates (z-scores) of all cognitive tasks (right). Middle: Network visualization (spring layout, left side) of CALM cortical volume data (N=246). Nodes are grouped according to Walktrap algorithm results (see above). Centrality estimates (z-scores) of all cortical volume nodes (right). Bottom: Network visualization (spring layout, left side) of CALM fractional anisotropy data (N=165). Nodes are grouped according to Walktrap algorithm results *after* outlier uncinate fasciculus and forceps major data (N=1 for each region, see above) was removed. Centrality estimates (z-scores) of all fractional anisotropy nodes (right). Dashed lines indicate mean strength and one standard deviation above the mean.

Bridging the gap: Multilayer networks

The regularized partial correlation network analyses for the CALM multilayer networks data are shown in Figure 5. Consistent with the pattern found in the single-layer networks, the cognitive and grey matter multilayer network (top left of Figure 5) edges are mostly positive and small to moderate weights (mean PC= 0.04, median PC=0, PC range= -0.12—0.64, Q from Walktrap=0.51). Comparably, the cognitive and white matter multilayer network (Figure 5, top right) had similar edge weight estimates (mean PC=0.04, median=0, range= -0.2—0.65, Q from Walktrap=0.41). Finally, combining all measures together (tri–layer network consisting of cognition, grey and white matter, bottom center of Figure 5) produced a network with similar characteristics to the bi-layer networks (mean PC=0.02, median PC=0, PC range= -0.2—0.66) except for modularity (Q=0.06). For the bi-layer networks, the Walktrap algorithm produced two clusters that consisted entirely of cognitive or neural nodes except for following instructions (Ins), which was either kicked out (cognition-grey matter) or grouped with a neural node (cognition-white matter, forceps minor). The result for the tri-layer networks was more complex with a total of 16 communities. However, one community (light blue) emerged as the largest, consisting of cognitive (mostly working memory), grey and white matter nodes (Figure 5, bottom center).

Regarding centrality, we report bridge (strength) centrality (Figure 6) to compare the results to the Walktrap algorithm. In the cognitive-grey matter network, four bridge nodes surfaced (in descending order: numerical operations, rostral middle frontal gyrus, digit recall and medial orbital frontal cortex, Figure 6 top left). In terms of stability, the CS-coefficient was 0, indicating that the bridge strength estimates were highly unstable under bootstrapping conditions (however, it must be noted that the age node had to be removed from data before estimation to enable calculation). In the cognitive-white matter bi-layer network, four nodes (in descending order: reading, inferior frontal-occipital fasciculus, spelling and peabody picture vocabulary test) emerged as possible bridge nodes (Figure 6, top right). Moreover, the centrality estimates had a CS-coefficient of 0.05, once again suggesting that the bridge strength estimates were unstable (although again the age node had to be removed from the data before estimation to enable calculation). Lastly, for the tri-layer network, 7 nodes displayed positive bridge strength equal to or greater than one standard deviation above the mean (Figure 6, bottom center). These included (in descending order): reading, superior temporal gyrus, forceps minor, corticospinal tract, spelling, anterior thalamic radiations, and peabody picture vocabulary test. Much better than the bi-layer networks, the tri-layer network bridge strength estimates were moderately stable (age was included in estimation, CS-coefficient=0.44).

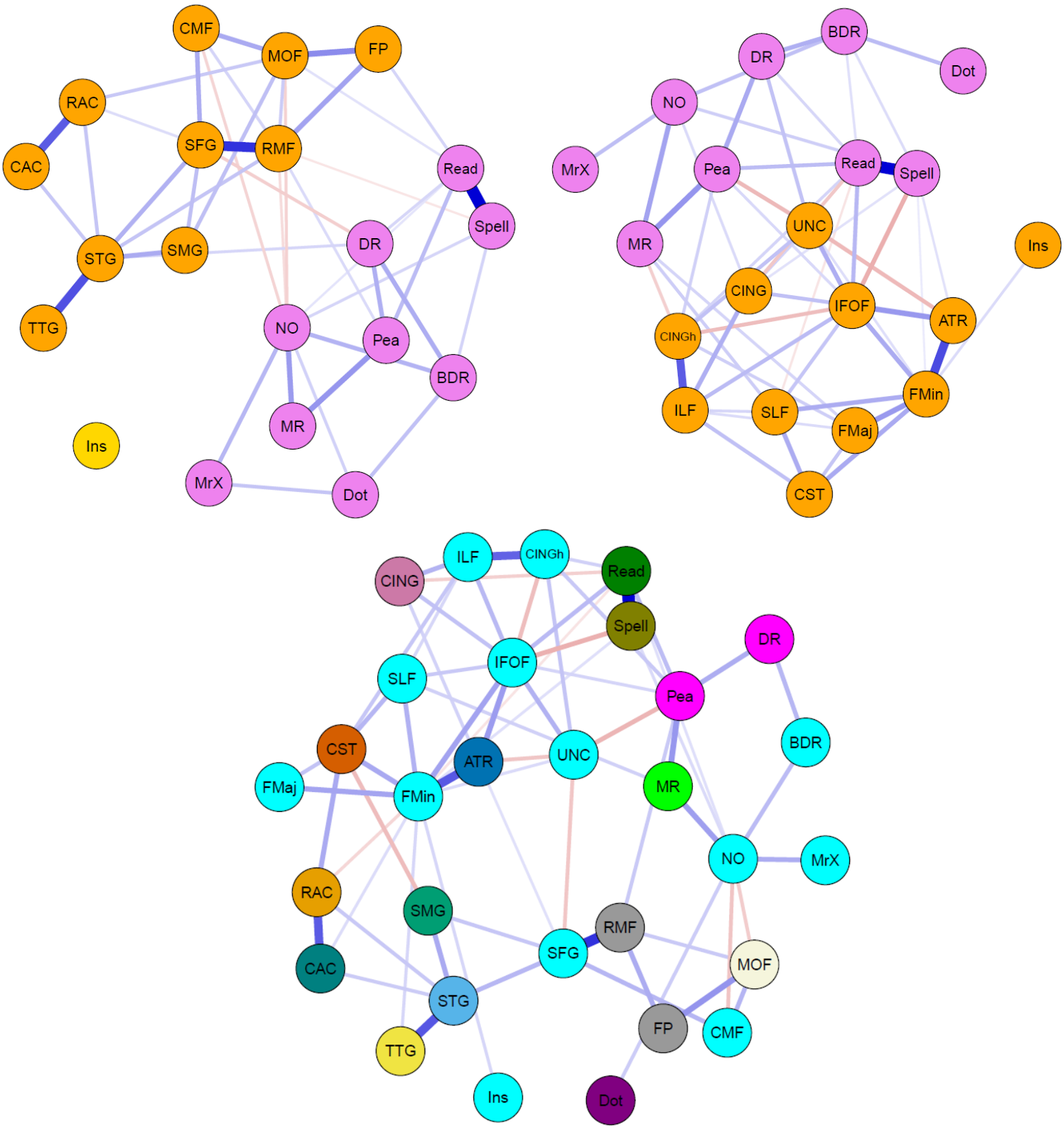
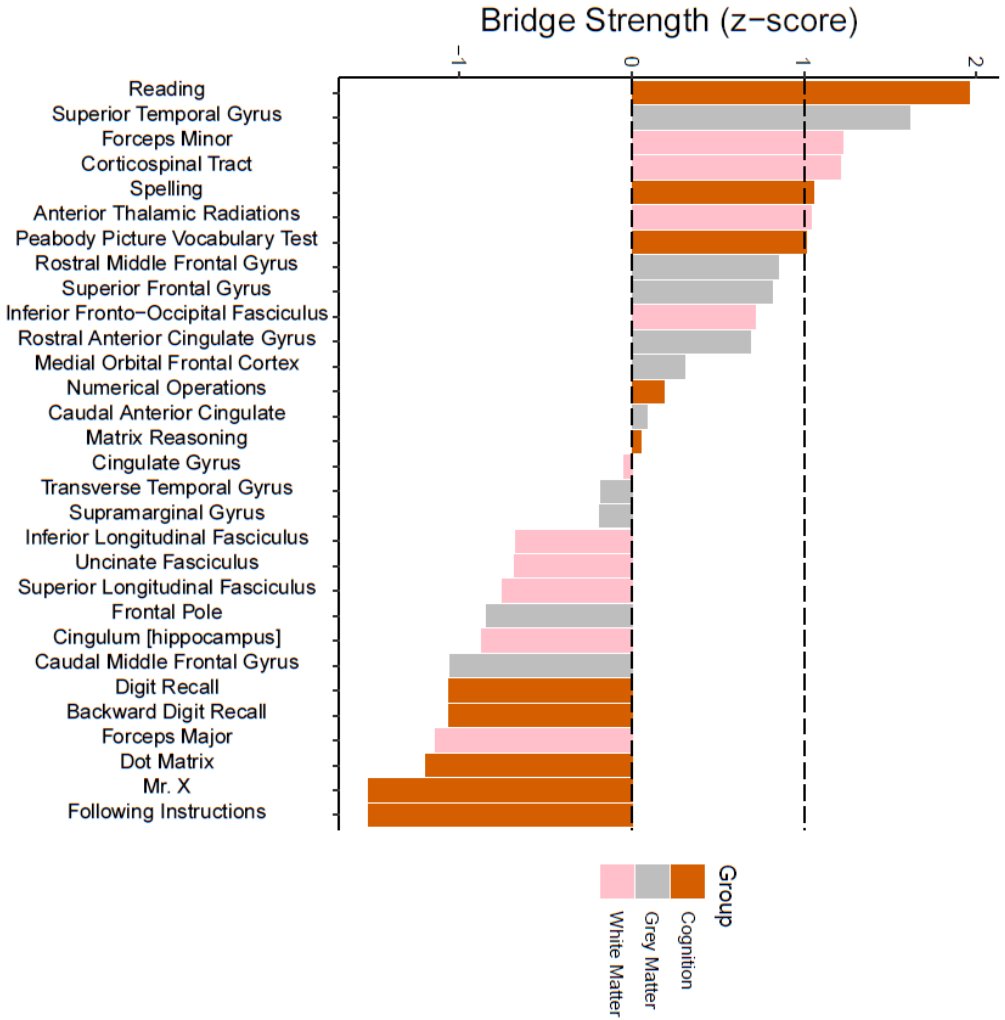


Figure 5. Network visualizations (spring layout) of partial correlation multilayer networks for CALM data. Nodes are grouped according to Walktrap algorithm results (see above).Top: Bi-layer networks consisting of cognition and grey matter (top left), and cognition and white matter (top right). Bottom: Tri-layer network consisting of cognition, grey matter and white matter (center).


Bridge Strength (z−score)
−1
0
1
2
Reading
Superior Temporal Gyrus
Forceps Minor
Corticospinal Tract
Spelling
Anterior Thalamic Radiations
Peabody Picture Vocabulary Test
Rostral Middle Frontal Gyrus
Superior Frontal Gyrus
Inferior Fronto−Occipital Fasciculus
Rostral Anterior Cingulate Gyrus
Medial Orbital Frontal Cortex
Numerical Operations
Caudal Anterior Cingulate
Matrix Reasoning
Cingulate Gyrus
Transverse Temporal Gyrus
Supramarginal Gyrus
Inferior Longitudinal Fasciculus
Uncinate Fasciculus
Superior Longitudinal Fasciculus
Frontal Pole
Cingulum [hippocampus]
Caudal Middle Frontal Gyrus
Digit Recall
Backward Digit Recall
Forceps Major
Dot Matrix
Mr. X
Following Instructions
Group
Cognition
Grey Matter
White Matter

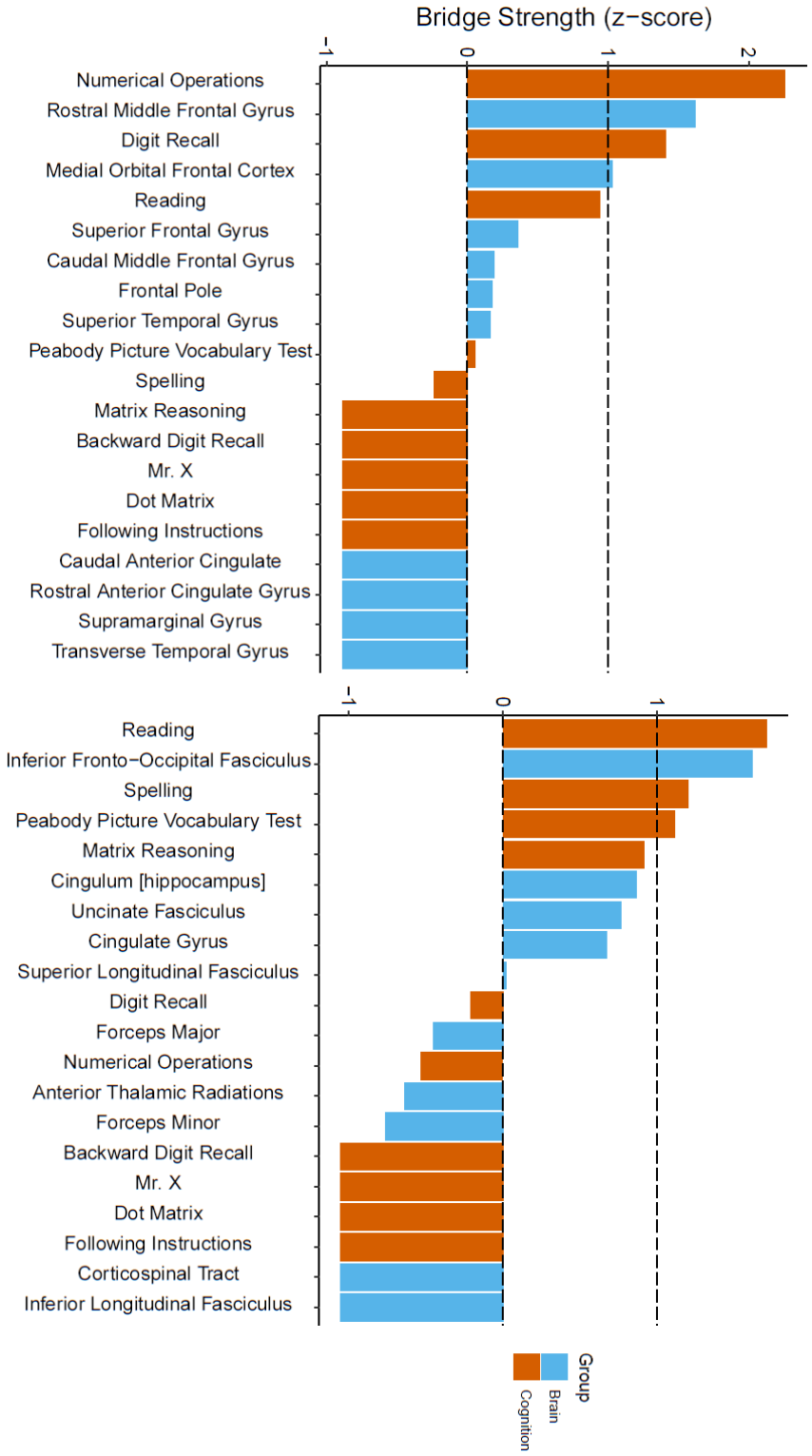

Bridge Strength (z−score)
−1
0
1
2
Numerical Operations
Rostral Middle Frontal Gyrus
Digit Recall
Medial Orbital Frontal Cortex
Reading
Superior Frontal Gyrus
Caudal Middle Frontal Gyrus
Frontal Pole
Superior Temporal Gyrus
Peabody Picture Vocabulary Test
Spelling
Matrix Reasoning
Backward Digit Recall
Mr. X
Dot Matrix
Following Instructions
Caudal Anterior Cingulate
Rostral Anterior Cingulate Gyrus
Supramarginal Gyrus
Transverse Temporal Gyrus
Reading
Inferior Fronto−Occipital Fasciculus
Spelling
Peabody Picture Vocabulary Test
Matrix Reasoning
Cingulum [hippocampus]
Uncinate Fasciculus
Cingulate Gyrus
Superior Longitudinal Fasciculus
Digit Recall
Forceps Major
Numerical Operations
Anterior Thalamic Radiations
Forceps Minor
Backward Digit Recall
Mr. X
Dot Matrix
Following Instructions
Corticospinal Tract
Inferior Longitudinal Fasciculus
Group
Brain
Cognition

Figure 6. Bridge centrality estimates (z-scores) for multilayer networks. Top: Bi-layer networks consisting of cognition and grey matter (top left), and cognition and white matter (top right). Bottom: Tri-layer network consisting of cognition, grey matter and white matter (center). Dashed lines indicate mean strength and one standard deviation above the mean.

For our final analysis, we applied the Clique Percolation algorithm (e.g. used for discovery of overlapping nodes) for weighted networks to our multilayer networks. In the cognitive-grey matter network, this produced four communities and 7 isolated nodes, which were mostly cognitive variables (Figure 7, top left). For the most part, these communities consisted of nodes that would typically be grouped together (behavior with behavior, brain with brain). However, two brain nodes (superior temporal gyrus and rostral middle frontal gyrus) shared membership between two communities. Both ROIs shared membership with the cognitive node digit recall, an assessment of working memory (Table 1). Moreover, all the other working memory tasks were classified as isolated nodes (no community membership). For the cognitive-white matter network, there were a total of four communities and 8 isolated nodes, though the result was more complex (Figure 7, top right). Most strikingly, two white matter nodes (inferior fronto-occipital fasciculus (here abbreviated as IFO) and cingulate gyrus (CING)) obtained membership to three communities. Lastly, for the tri-layer network (cognition, grey and white matter), the algorithm yielded five communities and 17 isolated nodes (Figure 7, bottom center). Only two nodes (superior temporal gyrus and inferior fronto-occipital fasciculus) shared memberships with other nodes, both of which were also found to show overlap with other nodes in the cognitive-grey matter and cognitive-white matter networks.

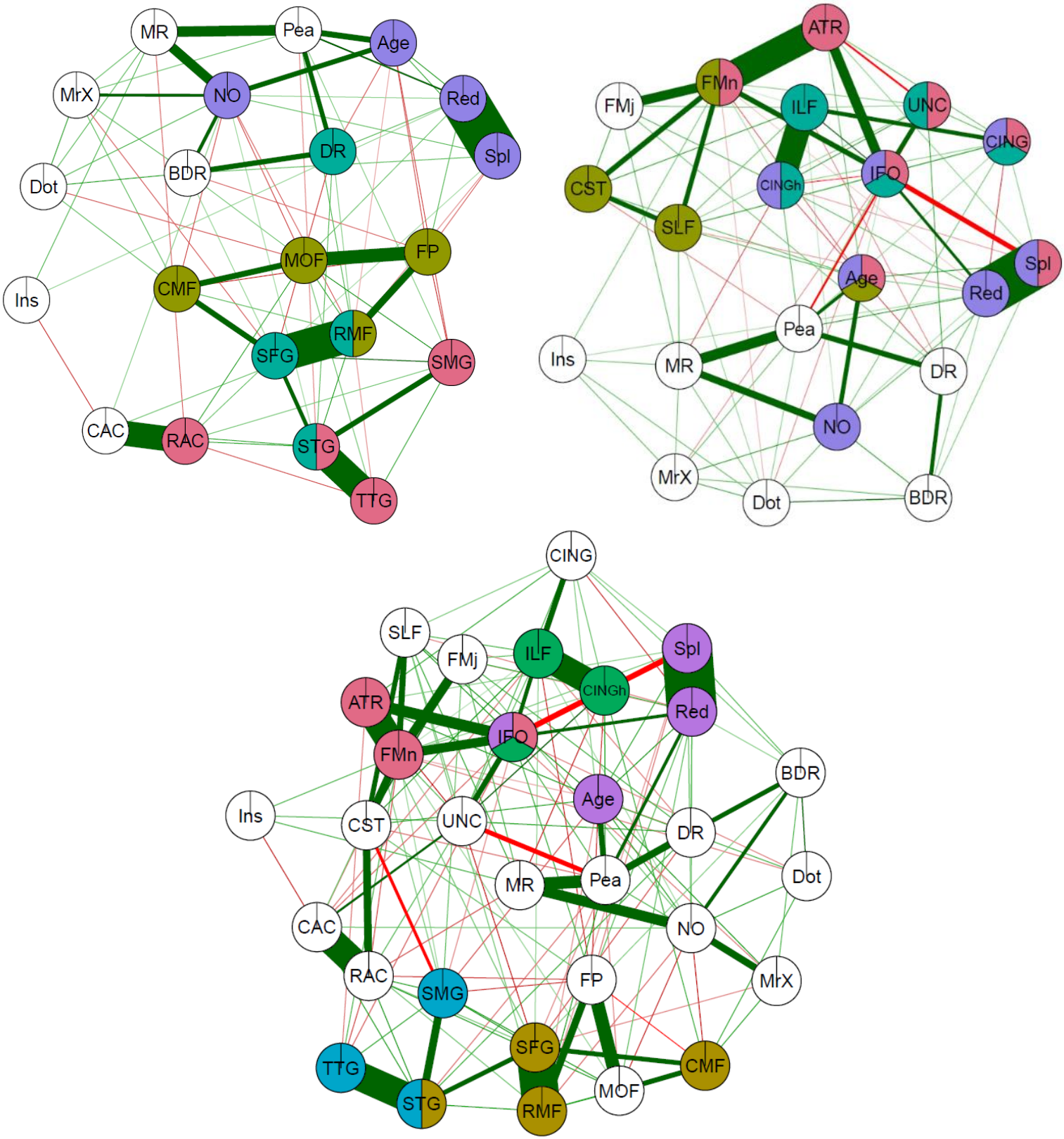


Figure 7. Community detection for cognition-grey matter (top left), cognition-white matter (top right) and tri-layer (cognition, grey matter and white matter) networks calculated by the Clique Percolation algorithm for small, weighted networks. White-filled nodes indicate isolated nodes. Note that the 'Age' node is shown in these networks but not in the previous ones, although the estimation procedure was the same.

DISCUSSION

Summary of main findings

In this study, we used network analysis (partial correlations) to examine the neurocognitive structure of general intelligence in a childhood and adolescent cohort of struggling learners (CALM). For our single-layer networks (Figure 4), we found that, within domains, cognitive, grey matter and white matter networks contained mostly (if not all) positive partial correlations. Moreover, in all single-layer networks, a subset of nodes emerged as more central than others (as indexed by node strength equal to or greater than one standard deviation above the mean), which varied in stability from questionable to highly reliable. In the cognitive network, this included verbal abilities (specifically reading). In the structural brain networks (grey matter cortical volume and white matter fractional anisotropy), three nodes emerged as central for the grey matter network (grey matter: (superior temporal gyrus, rostral middle frontal gyrus and superior frontal gyrus) while two white matter network nodes (forceps minor and inferior longitudinal fasciculus) passed as centrality threshold. Furthermore, we extended previous approaches by integrating networks of structural brain data with a cognitive network, forming bi- and tri-layer networks (Figure 5). Doing so, we observed multiple (both positive and negative) partial correlations between brain and behavior variables. Using bridge strength as a metric, we found that a mixture of cognitive and neural nodes harbor significant connections across communities (defined by the Walktrap algorithm) and levels of organization (Figure 6). Lastly, we used Clique Percolation analysis to identify nodes that overlap between various community memberships (behavior vs brain) and discovered a small group of nodes (e.g. inferior fronto-occipital fasciculus, rostral middle frontal gyrus, spelling and superior temporal gyrus, Figure 7) that corroborated some findings from bridge strength centrality. Overall, our results suggest which behavioral and neural variables have greater (possible) influence among or might be more influenced by other nodes and potentially serve as bridges

between the brain and cognition within general intelligence. However, the literature on drawing inferences from networks to the most likely consequences of intervening on the network is complex and rapidly changing, (e.g. Dablander & Hinne, 2019; Henry et al., 2020; Levine & Leucht, 2016).

Interpretation of network models and community detection analyses

For the cognitive network, each node corresponded to a single cognitive task (e.g. matrix reasoning) while partial correlations (weighted edges) between nodes were interpreted as compatible with (possible) causal consequences of interactions among cognitive abilities during development. This interpretation is compatible with the mutualism theory of cognitive development (Van Der Maas et al., 2006), whereby cognitive abilities positively reinforce each other (e.g. positive partial correlations) over time to produce the positive manifold (Spearman, 1904). Mutualism hypothesizes that general intelligence emerges from causal interactions among abilities rather than a general latent factor (Fried, 2020; Kan et al., 2019). Hence, cognition is viewed as a complex system derived from the dynamic relations of specific abilities that become more intertwined over development.

The existence of only positive edges in our cognitive network would be expected under a mutualistic perspective (interactions among cognitive variables), which at its essence is a network theory of general intelligence, although longitudinal analyses are need to further substantiate this claim. Initially, it was surprising that three of the four most central nodes (although reading was the only node with a positive strength value at or above one standard deviation) relate to verbal ability rather than abilities such as fluid intelligence and working memory (matrix reasoning and (forward and backward) digit recall), which are traditionally viewed as causal influences on cognitive development (Cattell, 1971). However, an emerging body of literature suggests that verbal ability plays a crucial role in cognitive development (e.g.

between reading and working memory before 4th grade, Peng et al., 2018 and Zhang & Malatesha Joshi, 2020, as well as driving the emergence of reasoning (Kievit et al., 2019; also see Gathercole et al., 1999).

As for our neural networks (here, grey matter cortical volume and white matter fractional anisotropy), individual nodes were comprised of a single ROIs. Importantly, we did not interpret weighted edges as an index of direct connectivity. Instead, the presence of strong associations between these ROIs would be compatible with the hypothesis of coordinated development (see Alexander-Bloch et al., 2013) whereby certain brain regions show preferential correlations to each other than more peripheral regions over time (e.g. childhood to late adolescence) as well as the notion of "rich" (Heuvel & Sporns, 2011) and "diverse" (Bertolero et al., 2017) clubs that enable local and global integration, respectively. Two of the three central cortical volume nodes (rostral middle frontal gyrus and superior frontal gyrus) display conserved spatial autocorrelation (are anatomically proximal to each other). However, the most central node, the superior temporal gyrus which has been implicated in verbal reasoning (e.g. Khundrakpam et al., 2017), remains distal from these regions. Regarding white matter, the two strongest nodes (forceps minor and inferior longitudinal fasciculus) while not anatomically close, instead represent long-range connections (see de Mooij et al., 2018) that have been linked to mathematical ability (Navas-Sánchez et al., 2014) and visuospatial working memory (Krogsrud et al., 2018).

We observed clusters in our cognitive network (defined by the Walktrap algorithm) that differ from the pre-assigned distinctions typically found in the literature (e.g. fluid intelligence and working memory). As with our grey matter nodes, both clusters correspond highly with their anatomical locations within the cortex, which falls in line with structural covariance hypothesis (Alexander-Bloch et al., 2013). In contrast, white matter clusters were not grouped according to anatomical proximity. However, given the size and reach (long-range cortical

connections) of these tracts, these results are still compatible with coordinated development and the pervasive findings of hubs in the case that they facilitate integration of information across spatially distributed areas of the brain, although longitudinal neurocognitive data and analysis are needed to make any strong claims about causal directionality.

Finally, we integrated both domains (cognitive abilities and brain metrics) into combined multilayer networks (cognition-grey matter, cognition-white matter and cognition-grey and- white matter). Doing so allowed us to attempt comparison and integration simultaneously across explanatory levels within the same analytical paradigm (network analysis) and statistical metric (partial correlations, centrality and community detection). From this analysis, three observations immediately stood out. First, there were multiple partial correlations between cognitive and neural nodes (especially in the cognitive-white matter and cognitive-grey matter and- white matter networks). Second, compared to the single-layer networks, the multilayer networks have more negative partial correlations. Together, these two findings further suggests that associations between the brain and cognition are complex as they defy straightforward (e.g. only positive and/or one-to-one) relationships and interpretations. However, it should be noted that causality (e.g. conditioning on colliders, see Rohrer, 2018 for overview of interpretations of correlations in graphical causal models in observational data) becomes even more difficult to determine with networks incorporating multiple levels of organization (e.g. cognition and structural brain covariance). Finally, we found a peculiar role of the cognitive task following instructions (Ins) within all multilayer networks. For example, in the cognitive-grey matter network, Ins had no partial correlations with any other nodes within the network while in both the cognitive-white matter and tri-layer network (cognition, grey and white matter) Ins *only* correlated with the forceps minor (FMin), a neural node, and not any of the cognitive variables. This might suggest that following instructions, traditionally a working memory task and often analyzed using structural equation modeling, may have

distinct psychometric properties (e.g. one-to-one mapping) when compared to other cognitive tasks when modeled through network science approaches, and/or when adjusted for all shared correlations.

Inspection of bridge strength centrality showed three interesting patterns. First, it appears that, on average (discounting the one standard deviation cutoff), the neural nodes are stronger than the cognitive variables within the multilayer networks, despite there being an equal number of cognitive nodes for each brain metric. This is possibly due to the large amount of edges between them (grey and white matter regions) and both cognitive and other neural nodes. In other words, since the neural nodes contain a larger number of connections (partial correlations) across explanatory levels, they display greater bridge strength (bridge strength sums inter-network correlations). Another convergent result between the multilayer networks was found in the working memory tasks. Backward digit recall, dot matrix, Mr. X and following instructions were consistently among the most negative nodes in terms of bridge strength centrality. This might suggest that, in struggling learners, working memory serves more as a bottleneck on baseline cognition rather than helping to facilitate the development general intelligence (e.g. mutualism). Moreover, in the tri-layer network, we observed a convergence of bridge nodes but only for the cognitive-white matter network whereby the central nodes in that network (specifically the peabody picture vocabulary test, reading and spelling) were also central in the tri-layer network.

In other ways, the multilayer networks differed. First, in the tri-layer network, the positive bridge strength estimates (including those classified as central) were more likely to be neural rather than cognitive nodes. Secondly, numerical operations exhibited high bridge strength only in the cognitive-grey matter network while displaying negative bridge strength in the cognitive-white matter network, suggesting that grey and white matter, while related, possibly play distinct roles in mathematical ability. Thirdly, for the tri-layer network, none of

the top three strongest neural nodes (in descending order: superior temporal gyrus, forceps minor and corticospinal tract) were also central in the bi-layer networks, which suggests that simultaneously analyzing grey and white matter together with cognition might provide unique information about brain-behavior relationships.

Finally, in terms of the Clique Percolation algorithm, the cognitive-white matter network contained several more two-membership nodes (e.g. hippocampal cingulum, spelling, etc.) compared to the cognitive-grey matter network whose nodes tended to stay within their respective domains (brain or behavior). Only two nodes (grey matter: superior temporal gyrus; white matter: inferior fronto-occipital fasciculus) shared overlapping membership both in their own respective networks and the tri-layer network. Interestingly, only the superior temporal gyrus was found to be central in the tri-layer network while the inferior fronto-occipital fasciculus was the second most central node in the cognitive-white matter bi-layer network. The superior temporal gyrus has been specifically implicated in verbal reasoning (Khundrakpam et al., 2017), corresponding to our finding that verbal ability plays a central role in general intelligence, but has also been linked to aspects of cognition such as executive (dys)function, fluid intelligence and working memory (Grambaite et al., 2011 and Pineda-Pardo et al., 2016, respectively, although using white matter diffusivity and fractional anisotropy). For the tri-layer network, only the IFOF shared membership with behavioral tasks, specifically verbal abilities (reading and spelling), which were also stronger in our single-layer cognitive network. Together, these findings somewhat contrast with our bridge strength results.

Limitations of the current study

This study contains several limitations that require caution when interpreting the results. First and foremost, these findings are based on cross-sectional data. While adequate to help tease apart individual differences in cognition *between* people, cross-sectional data cannot

be used to elucidate differences in *changes within* individuals over time, such as during development. Therefore, longitudinal analyses are needed before attempting to make strong inferences about the dynamics of these networks. Moreover, the CALM sample represents an atypical sample (Holmes et al., 2019), with participants who consistently score lower on measures of attention, learning and/or memory than age-matched controls (see Figure 2 (Level I) of Simpson-Kent et al., 2020 for comparison to a typically developing sample). As a result, these analyses would need to be replicated in additional (ideally larger) samples with different cognitive profiles before our results can be generalized. Last, but not least, as seen with white matter covariance, our sample contained an outlier that affected the Walktrap algorithm despite rigorous quality control protocols. Therefore, we re-ran our analyses to test the sensitivity of our main findings (e.g. positive partial correlations and central nodes) to potential outliers (defined as ± 4 standard deviations). Doing so did not severely alter our findings except for the white matter covariance network mentioned above and the cognitive-grey matter network, which showed much less interconnectivity between brain and behavior, suggesting that bridge centrality was biased due to an overrepresentation of strong intra-correlations among cognitive and neural nodes. However, it must be restated that our data comes from an atypical sample, which might influence brain metrics even with rigorous quality control procedures. Therefore, despite this discrepancy, our data supports brain-behavior 'bridges' in general intelligence, especially in white matter microstructure.

Future directions toward theory building in cognitive neuroscience

Our results that suggest verbal abilities rather than fluid intelligence or working memory might play a more pivotal role in the development of cognitive ability fits with the gradual progression in schooling. For example, before children can successfully be taught more advanced subjects (e.g. history, reading comprehension, etc.), they must first become

competent in basic language faculties. In other words, it may be that verbal skills (e.g. reading and spelling) facilitate performance on abstract tests, even in the absence of direct knowledge-based task demands. Recent evidence has been found supporting this notion and suggest that verbal ability, particularly reading and vocabulary in relation to working memory and reasoning, might drive early cognitive development (Kievit et al., 2019; Peng et al., 2018; Zhang & Malatesha Joshi, 2020). Therefore, future studies could further examine whether greater verbal ability in early development facilitates greater acquisition of higher-level cognitive skills by lowering computational demands in working memory.

Moreover, in this context, the fact that the numerical operations task was also found to be central (cognitive-grey matter network only) should be expected since mathematics (e.g. arithmetic) also involves symbol manipulation. In terms of mutualism (Van Der Maas et al., 2006), future models (ideally in longitudinal samples) could test whether language and other symbolic abilities show progressively higher reciprocal associations during early development compared to other abilities until more complex cognition (i.e. fluid reasoning and working memory) develops in later childhood (also see Kievit et al., 2019 and Peng et al., 2018).

Next, concerning the centrality and clustering of nodes found in our brain structural covariance networks, we hypothesize that, whether anatomically close to each other (grey matter) or forming long-range connections throughout the cortex, especially frontal, parietal and temporal areas (white matter), central and highly clustered regions enable both “easy-to-reach” (knowledge for specific abilities acquired through experience) and “difficult-to-reach” (necessary for integrating information for complex problem-solving) network states (Barbey, 2018). However, this hypothesis needs further investigating as current evidence regarding Barbey’s theory, while encouraging, remains inconclusive (Girn et al., 2019). This would need to be tested in both longitudinal and cross-species cohorts, which is outside the scope of this study.

Finally, our discoveries using the weighted Clique Percolation algorithm (Farkas et al., 2007) further emphasized the 'bridge' concept of nodes, especially neural nodes (e.g. grey matter: rostral middle frontal gyrus and superior temporal gyrus; white matter: hippocampal cingulum, cingulate gyrus and inferior frontal-occipital fasciculus), existing and functioning between the brain and cognition. Therefore, these results suggest that future (longitudinal) studies could manipulate (e.g. transcranial magnetic stimulation) or study disruptions to these regions (e.g. traumatic brain injury, tumors, lesions, etc.) and measure change in verbal skills such as reading and spelling.

Fortunately, two recent studies have taken on this challenge. Bertolero et al., 2018 found that a mechanistic model assuming that "connector hubs" (diverse club nodes, see Bertolero et al., 2017), which regulate the activity of their neighboring communities to be more modular but maintain the capability of "task appropriate information integration across communities", significantly predicted higher cognitive performance on various tasks including language and working memory. Furthermore, in the same sample studied here, Akarca et al., 2020 applied a generative network modelling approach to simulate the growth of brain network connectomes, finding that it is possible to simulate structural networks with statistical properties mirroring the spatially embedding of those observed. The parameters of these generative models were shown to correlate with neuroimaging measures not used to train the models (including grey matter measures), cognitive performance (including vocabulary and mathematics) and relate to gene expression in the cortex. Together these studies point the field toward a better mechanistic understanding of the development of human brain structure, function and their relationship with cognitive ability.

DECLARATIONS OF INTEREST

EB is a member of the scientific advisory board of Sosei Heptares.

ACKNOWLEDGMENTS

The Centre for Attention Learning and Memory (CALM) research clinic is based at and supported by funding from the MRC Cognition and Brain Sciences Unit, University of Cambridge. The Principal Investigators are Joni Holmes (Head of CALM), Susan Gathercole (Chair of CALM Management Committee), Duncan Astle, Tom Manly, Kate Baker and Rogier Kievit. Data collection is assisted by a team of researchers and PhD students at the CBU that includes Joe Bathelt, Giacomo Bignardi, Sarah Bishop, Erica Bottacin, Lara Bridge, Annie Bryant, Sally Butterfield, Elizabeth Byrne, Gemma Crickmore, Edwin Dalmaijer, Fanchea Daily, Tina Emery, Laura Forde, Grace Franckel, Delia Fuhrmann, Andrew Gadie, Sara Gharooni, Jacalyn Guy, Erin Hawkins, Agniezska Jaroslawska, Amy Johnson, Jonathon Jones, Silvana Mareva, Elise Ng-Cordell, Sinead O'Brien, Cliodhna O'Leary, Joseph Rennie, Ivan Simpson-Kent, Roma Siugzdaite, Tess Smith, Stepheni Uh, Francesca Woolgar and Mengya Zhang. The authors wish to thank the many professionals working in children's services in the South-East and East of England for their support, and to the children and their families for giving up their time to visit the clinic.

I.L.S.-K. is supported by the Cambridge Trust., E.I.F. did not receive funding for this project, S.M. by the Medical Research Council PhD Studentship and D.A. by the Medical Research Council Doctoral Training Partnership Studentship (Project Code: SUAH/012; Award: RG86932) and the Cambridge Trust. E.T.B. is supported by an NIHR Senior Investigator award. This project also received funding from the European Union's Horizon 2020 research and innovation programme (grant agreement number 732592). R.A.K. is supported by the Wellcome Trust (Grant No. 107392/Z/15/Z), the UK Medical Research Council SUAG/014 RG91365 and the Hypatia Fellowship by Radboud University.

REFERENCES

Akarca, D., Vértes, P. E., Bullmore, E. T., Team, the C., & Astle, D. E. (2020). A generative network model of neurodevelopment. *BioRxiv*, 2020.08.13.249391. https://doi.org/10.1101/2020.08.13.249391

Alexander-Bloch, A., Giedd, J. N., & Bullmore, E. (2013). Imaging structural co-variance between human brain regions. *Nature Reviews Neuroscience*, *14*(5), 322–336. https://doi.org/10.1038/nrn3465

Alloway, T. P. (2007). *Automated Working Memory Assessment (AWMA)*. London: Harcourt Assessment.

Amestoy et al., S. A. file igraph author. (2020). *igraph: Network Analysis and Visualization* (1.2.5) [Computer software]. https://CRAN.R-project.org/package=igraph

Assem, M., Glasser, M. F., Van Essen, D. C., & Duncan, J. (2020). A Domain-General Cognitive Core Defined in Multimodally Parcellated Human Cortex. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhaa023

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Barabási, A.-L. (2016). *Network Science*. Cambridge University Press. http://networksciencebook.com/

Barbey, A. K. (2018). Network Neuroscience Theory of Human Intelligence. *Trends in Cognitive Sciences*, *22*(1), 8–20. https://doi.org/10.1016/j.tics.2017.10.001

Bassett, Danielle S., & Bullmore, E. T. (2017). Small-World Brain Networks Revisited. *The Neuroscientist*, *23*(5), 499–516. https://doi.org/10.1177/1073858416667720

Bassett, Danielle S, & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*(3), 353–364. https://doi.org/10.1038/nn.4502

Bassett, Danielle Smith, & Bullmore, E. (2006). Small-World Brain Networks. *The Neuroscientist*, *12*(6), 512–523. https://doi.org/10.1177/1073858406293182

Basten, U., Hilger, K., & Fiebach, C. J. (2015). Where smart brains are different: A quantitative meta-analysis of functional and structural brain imaging studies on intelligence. *Intelligence*, *51*, 10–27. https://doi.org/10.1016/j.intell.2015.04.009

Bertolero, M. A., Yeo, B. T. T., & D'Esposito, M. (2017). The diverse club. *Nature Communications*, *8*(1), 1277. https://doi.org/10.1038/s41467-017-01189-w

Bertolero, Maxwell A., Yeo, B. T. T., Bassett, D. S., & D'Esposito, M. (2018). A mechanistic model of connector hubs, modularity and cognition. *Nature Human Behaviour*, *2*(10), 765–777. https://doi.org/10.1038/s41562-018-0420-6

Bianconi, G. (2018). *Multilayer Networks: Structure and Function*. Oxford University Press.

Borkulo, C. V., Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R., Borsboom, D., & Waldorp, L. (2017). Comparing network structures on three aspects: A permutation test. *Unpublished*. https://doi.org/10.13140/rg.2.2.29455.38569

Borkulo, C. van, Epskamp, S., Jones, P., Haslbeck, J., & Millner, A. (2019). *NetworkComparisonTest: Statistical Comparison of Two Networks Based on Three Invariance Measures* (2.2.1) [Computer software]. https://CRAN.R-project.org/package=NetworkComparisonTest

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*(5), 336–349. https://doi.org/10.1038/nrn3214

Calvin, C. M., Deary, I. J., Fenton, C., Roberts, B. A., Der, G., Leckenby, N., & Batty, G. D. (2011). Intelligence in youth and all-cause-mortality: Systematic review with meta-analysis. *International Journal of Epidemiology*, *40*(3), 626–644. https://doi.org/10.1093/ije/dyq190

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action* (pp. xxii, 583). Houghton Mifflin.

Chuderski, A. (2015). The broad factor of working memory is virtually isomorphic to fluid intelligence tested under time pressure. *Personality and Individual Differences*, *85*, 98–104. https://doi.org/10.1016/j.paid.2015.04.046

Dablander, F., & Hinne, M. (2019). Node centrality measures are a poor substitute for causal inference. *Scientific Reports*, *9*(1), 1–13. https://doi.org/10.1038/s41598-019-43033-9

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

de Mooij, S. M. M., Henson, R. N. A., Waldorp, L. J., & Kievit, R. A. (2018). Age Differentiation within Gray Matter, White Matter, and between Memory and White Matter in an Adult Life Span Cohort. *The Journal of Neuroscience*, *38*(25), 5826–5836. https://doi.org/10.1523/JNEUROSCI.1627-17.2018

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The Impact of Childhood Intelligence on Later Life: Following Up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology*, *86*(1), 130–147. https://doi.org/10.1037/0022-3514.86.1.130

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test*. Pearson Assessments.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Costantini, G., Haslbeck, J., Isvoranu, A., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2020). *qgraph: Graph Plotting Methods, Psychometric Data Visualization and Graphical Model Estimation* (1.6.5) [Computer software]. https://CRAN.R-project.org/package=qgraph

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167

Epskamp, S., & Fried, E. I. (2020). *bootnet: Bootstrap Methods for Various Network Estimation Routines* (1.4.2) [Computer software]. https://CRAN.R-project.org/package=bootnet

Epskamp, S., Maris, G., Waldorp, L. J., & Borsboom, D. (2018). Network Psychometrics. In *The Wiley Handbook of Psychometric Testing* (pp. 953–986). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118489772.ch30

Farkas, I., Ábel, D., Palla, G., & Vicsek, T. (2007). Weighted network modules. *New Journal of Physics*, *9*(6), 180–180. https://doi.org/10.1088/1367-2630/9/6/180

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*(2), 195–207. https://doi.org/10.1006/nimg.1998.0396

Fischl, Bruce, & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, *97*(20), 11050–11055. https://doi.org/10.1073/pnas.200033797

Fischl, Bruce, Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuron*, *33*(3), 341–355. https://doi.org/10.1016/S0896-6273(02)00569-X

Fornito, A., & Bullmore, E. T. (2012). Connectomic Intermediate Phenotypes for Psychiatric Disorders. *Frontiers in Psychiatry*, *3*. https://doi.org/10.3389/fpsyt.2012.00032

Fornito, A., Zalesky, A., & Bullmore, E. (2016). *Fundamentals of Brain Network Analysis*. Academic Press.

Fried, E. I. (2020). *Lack of theory building and testing impedes progress in the factor and network literature*. https://doi.org/10.17605/OSF.IO/6CTS9

Fried, E. I., & Robinaugh, D. J. (2020). Systems all the way down: Embracing complexity in mental health research. *BMC Medicine*, *18*(1), 205. https://doi.org/10.1186/s12916-020-01668-w

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441. https://doi.org/10.1093/biostatistics/kxm045

Gates, K. M., Henry, T., Steinley, D., & Fair, D. A. (2016). A Monte Carlo Evaluation of Weighted Community Detection Algorithms. *Frontiers in Neuroinformatics*, *10*. https://doi.org/10.3389/fninf.2016.00045

Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology*, *22*(8), 1019–1037. https://doi.org/10.1002/acp.1407

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A.-M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology*, *13*(1), 65–77. https://doi.org/10.1002/(SICI)1099-0720(199902)13:1<65::AID-ACP548>3.0.CO;2-O

Girn, M., Mills, C., & Christoff, K. (2019). Linking brain network reconfiguration and intelligence: Are we there yet? *Trends in Neuroscience and Education*, *15*, 62–70. https://doi.org/10.1016/j.tine.2019.04.001

Graham, M. S., Drobnjak, I., & Zhang, H. (2016). Realistic simulation of artefacts in diffusion MRI for validating post-processing correction techniques. *NeuroImage*, *125*, 1079–1094. https://doi.org/10.1016/j.neuroimage.2015.11.006

Grambaite, R., Selnes, P., Reinvang, I., Aarsland, D., Hessen, E., Gjerstad, L., & Fladby, T. (2011). Executive Dysfunction in Mild Cognitive Impairment is Associated with

Changes in Frontal and Cingulate White Matter Tracts. *Journal of Alzheimer's Disease*, *27*(2), 453–462. https://doi.org/10.3233/JAD-2011-110290

Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q. K., Yu, A. B., Kahn, A. E., Medaglia, J. D., Vettel, J. M., Miller, M. B., Grafton, S. T., & Bassett, D. S. (2015). Controllability of structural brain networks. *Nature Communications*, *6*(1), 1–10. https://doi.org/10.1038/ncomms9414

Hegelund, E. R., Flensborg-Madsen, T., Dammeyer, J., & Mortensen, E. L. (2018). Low IQ as a predictor of unsuccessful educational and occupational achievement: A register-based study of 1,098,742 men in Denmark 1968–2016. *Intelligence*, *71*, 46–53. https://doi.org/10.1016/j.intell.2018.10.002

Henry, T. R., Robinaugh, D., & Fried, E. I. (2020). *On the control of psychological networks*. PsyArXiv. https://doi.org/10.31234/osf.io/7vpz2

Heuvel, M. P. van den, & Sporns, O. (2011). Rich-Club Organization of the Human Connectome. *Journal of Neuroscience*, *31*(44), 15775–15786. https://doi.org/10.1523/JNEUROSCI.3539-11.2011

Hilland, E., Landrø, N. I., Kraft, B., Tamnes, C. K., Fried, E. I., Maglanoc, L. A., & Jonassen, R. (2020). Exploring the links between specific depression symptoms and brain structure: A network study. *Psychiatry and Clinical Neurosciences*, *74*(3), 220–221. https://doi.org/10.1111/pcn.12969

Holmes, J., Bryant, A., Gathercole, S. E., & the CALM Team. (2019). Protocol for a transdiagnostic study of children with problems of attention, learning and memory (CALM). *BMC Pediatrics*, *19*(10). https://doi.org/10.1186/s12887-018-1385-3

Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D. S., Calabresi, P. A., Pekar, J. J., van Zijl, P. C. M., & Mori, S. (2008). Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. *NeuroImage*, *39*(1), 336–347. https://doi.org/10.1016/j.neuroimage.2007.07.053

Jones, P. (2020). *networktools: Tools for Identifying Important Nodes in Networks* (1.2.3) [Computer software]. https://CRAN.R-project.org/package=networktools

Jones, P. J., Ma, R., & McNally, R. J. (2019). Bridge Centrality: A Network Approach to Understanding Comorbidity. *Multivariate Behavioral Research*, *0*(0), 1–15. https://doi.org/10.1080/00273171.2019.1614898

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*(02), 135. https://doi.org/10.1017/S0140525X07001185

Kan, K.-J., van der Maas, H. L. J., & Levine, S. Z. (2019). Extending psychometric network analysis: Empirical evidence against g in favor of mutualism? *Intelligence*, *73*, 52–62. https://doi.org/10.1016/j.intell.2018.12.004

Kenett, D. Y., Perc, M., & Boccaletti, S. (2015). Networks of networks – An introduction. *Chaos, Solitons & Fractals*, *80*, 1–6. https://doi.org/10.1016/j.chaos.2015.03.016

Khundrakpam, B. S., Lewis, J. D., Reid, A., Karama, S., Zhao, L., Chouinard-Decorte, F., & Evans, A. C. (2017). Imaging structural covariance in the development of intelligence. *NeuroImage*, *144*, 227–240. https://doi.org/10.1016/j.neuroimage.2016.08.041

Kievit, R. A., Davis, S. W., Griffiths, J., Correia, M. M., Cam-CAN, & Henson, R. N. (2016). A watershed model of individual differences in fluid intelligence. *Neuropsychologia*, *91*, 186–198. https://doi.org/10.1016/j.neuropsychologia.2016.08.008

Kievit, R. A., Fuhrmann, D., Borgeest, G. S., Simpson-Kent, I. L., & Henson, R. N. A. (2018). The neural determinants of age-related changes in fluid intelligence: A pre-registered, longitudinal analysis in UK Biobank. *Wellcome Open Research*, *3*. https://doi.org/10.12688/wellcomeopenres.14241.2

Kievit, R. A., Hofman, A. D., & Nation, K. (2019). Mutualistic Coupling Between Vocabulary and Reasoning in Young Children: A Replication and Extension of the Study by Kievit et al. (2017). *Psychological Science*, *30*(8), 1245–1252. https://doi.org/10.1177/0956797619841265

Kievit, R. A., Lindenberger, U., Goodyer, I. M., Jones, P. B., Fonagy, P., Bullmore, E. T., & Dolan, R. J. (2017). Mutualistic Coupling Between Vocabulary and Reasoning Supports Cognitive Development During Late Adolescence and Early Adulthood.

Psychological Science, *28*(10), 1419–1431. https://doi.org/10.1177/0956797617710785

Kievit, R., Frankenhuis, W. E., Waldorp, L., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00513

Kievit, R., & Simpson-Kent, I. L. (2019). *It's About Time: Towards a Longitudinal Cognitive Neuroscience of Intelligence* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/n2yg7

Krogsrud, S. K., Fjell, A. M., Tamnes, C. K., Grydeland, H., Due-Tønnessen, P., Bjørnerud, A., Sampaio-Baptista, C., Andersson, J., Johansen-Berg, H., & Walhovd, K. B. (2018). Development of white matter microstructure in relation to verbal and visuospatial working memory—A longitudinal study. *PLOS ONE*, *13*(4), e0195540. https://doi.org/10.1371/journal.pone.0195540

Lange, J., & Zickfeld, J. (2019). *CliquePercolation: Clique Percolation for Networks* (0.2.0) [Computer software]. https://CRAN.R-project.org/package=CliquePercolation

Levenstein, D., Alvarez, V. A., Amarasingham, A., Azab, H., Gerkin, R. C., Hasenstaub, A., Iyer, R., Jolivet, R. B., Marzen, S., Monaco, J. D., Prinz, A. A., Quraishi, S., Santamaria, F., Shivkumar, S., Singh, M. F., Stockton, D. B., Traub, R., Rotstein, H. G., Nadim, F., & Redish, A. D. (2020). On the role of theory and modeling in neuroscience. *ArXiv:2003.13825 [q-Bio]*. http://arxiv.org/abs/2003.13825

Levine, S. Z., & Leucht, S. (2016). Identifying a system of predominant negative symptoms: Network analysis of three randomized clinical trials. *Schizophrenia Research*, *178*(1), 17–22. https://doi.org/10.1016/j.schres.2016.09.002

Mareva, S., & Holmes, J. (2020). *Network models of learning and cognition in typical and atypical learners*. https://doi.org/10.31219/osf.io/3tn5m

Mareva, S., Holmes, J., & the CALM team. (2019). Transdiagnostic associations across communication, cognitive, and behavioural problems in a developmentally at-risk

population: A network approach. *BMC Pediatrics*, *19*(1), 452. https://doi.org/10.1186/s12887-019-1818-7

Meunier, D., Lambiotte, R., & Bullmore, E. T. (2010). Modular and Hierarchically Modular Organization of Brain Networks. *Frontiers in Neuroscience*, *4*. https://doi.org/10.3389/fnins.2010.00200

Navas-Sánchez, F. J., Alemán-Gómez, Y., Sánchez-Gonzalez, J., Guzmán-De-Villoria, J. A., Franco, C., Robles, O., Arango, C., & Desco, M. (2014). White matter microstructure correlates of mathematical giftedness and intelligence quotient: White Matter Microstructure. *Human Brain Mapping*, *35*(6), 2619–2631. https://doi.org/10.1002/hbm.22355

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103

Palla, G., Derényi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, *435*(7043), 814–818. https://doi.org/10.1038/nature03607

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., Dardick, W., & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin*, *144*(1), 48–76. https://doi.org/10.1037/bul0000124

Pineda-Pardo, J. A., Martínez, K., Román, F. J., & Colom, R. (2016). Structural efficiency within a parieto-frontal network and cognitive differences. *Intelligence*, *54*, 105–116. https://doi.org/10.1016/j.intell.2015.12.002

Pons, P., & Latapy, M. (2005). Computing Communities in Large Networks Using Random Walks. In pInar Yolum, T. Güngör, F. Gürgen, & C. Özturan (Eds.), *Computer and Information Sciences—ISCIS 2005* (pp. 284–293). Springer. https://doi.org/10.1007/11569596_31

R Core Team. (2020). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.* https://www.R-project.org

Rhemtulla, M., Bork, R. van, & Cramer, A. O. J. (2020). Cross-lagged network models. *Multivariate Behavioral Research*. https://research.tilburguniversity.edu/en/publications/cross-lagged-network-models

Robinaugh, D. J., Hoekstra, R. H. A., Toner, E. R., & Borsboom, D. (2019). The network approach to psychopathology: A review of the literature 2008–2018 and an agenda for future research. *Psychological Medicine*, 1–14. https://doi.org/10.1017/S0033291719003404

Rohrer, J. M. (2018). Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42. https://doi.org/10.1177/2515245917745629

Schmank, C. J., Goring, S. A., Kovacs, K., & Conway, A. R. A. (2019). Psychometric Network Analysis of the Hungarian WAIS. *Journal of Intelligence*, *7*(3), 21. https://doi.org/10.3390/jintelligence7030021

Seidlitz, J., Váša, F., Shinn, M., Romero-Garcia, R., Whitaker, K. J., Vértes, P. E., Wagstyl, K., Kirkpatrick Reardon, P., Clasen, L., Liu, S., Messinger, A., Leopold, D. A., Fonagy, P., Dolan, R. J., Jones, P. B., Goodyer, I. M., Raznahan, A., & Bullmore, E. T. (2018). Morphometric Similarity Networks Detect Microscale Cortical Organization and Predict Inter-Individual Cognitive Variation. *Neuron*, *97*(1), 231-247.e7. https://doi.org/10.1016/j.neuron.2017.11.039

Simpson-Kent, I. L., Fuhrmann, D., Bathelt, J., Achterberg, J., Borgeest, G. S., & Kievit, R. A. (2020). Neurocognitive reorganization between crystallized intelligence, fluid intelligence and white matter microstructure in two age-heterogeneous developmental cohorts. *Developmental Cognitive Neuroscience*, *41*, 100743. https://doi.org/10.1016/j.dcn.2019.100743

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–155. https://doi.org/10.1002/hbm.10062

Solé-Casals, J., Serra-Grabulosa, J. M., Romero-Garcia, R., Vilaseca, G., Adan, A., Vilaró, N., Bargalló, N., & Bullmore, E. T. (2019). Structural brain network of gifted children

has a more integrated and versatile topology. *Brain Structure and Function*, *224*(7), 2373–2383. https://doi.org/10.1007/s00429-019-01914-9

Spearman, C. (1904). “General Intelligence,” Objectively Determined and Measured. *The American Journal of Psychology*, *15*(2), 201–292.

Sporns, O., & Betzel, R. F. (2016). Modular Brain Networks. *Annual Review of Psychology*, *67*, 613–640. https://doi.org/10.1146/annurev-psych-122414-033634

Van Der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*(4), 842–861. https://doi.org/10.1037/0033-295X.113.4.842

Van Der Maas, H. L. J., Kan, K.-J., Marsman, M., & Stevenson, C. E. (2017). Network Models for Cognitive Development and Intelligence. *Journal of Intelligence*, *5*(2), 16. https://doi.org/10.3390/jintelligence5020016

Wandell, B. A. (2016). Clarifying Human White Matter. *Annual Review of Neuroscience*, *39*(1), 103–128. https://doi.org/10.1146/annurev-neuro-070815-013815

Wechsler, D. (2005). *Wechsler Individual Achievement Test-Second UK Edition (WIAT-II)*. Pearson: London, UK.

Wechsler, D. (2008). *Wechsler adult intelligence Scale–Fourth edition (WAIS–IV)*. The Psychological Corporation, San Antonio, TX.

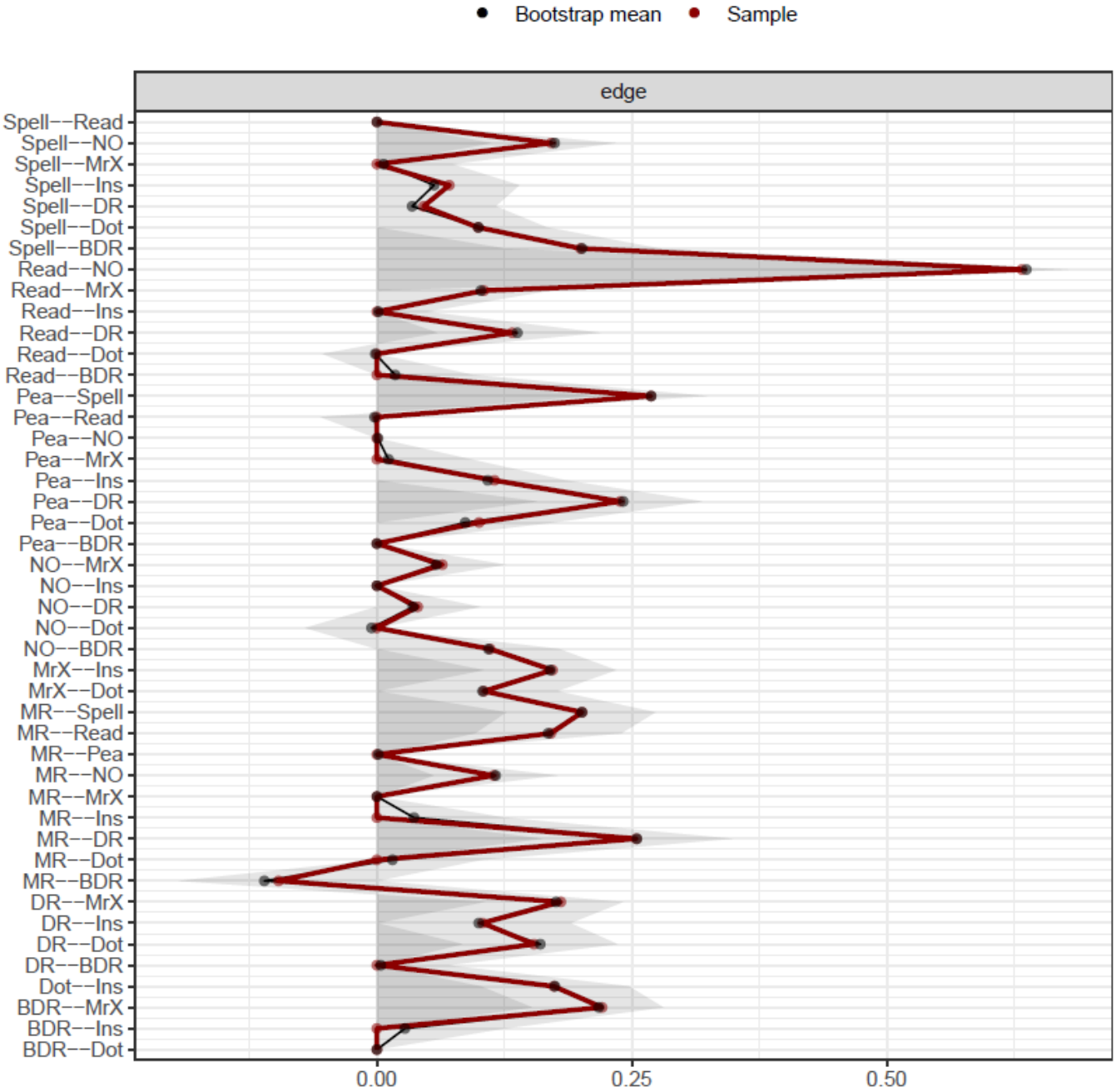
Wechsler, D. (2011). *Wechsler Abbreviated Scales of Intelligence-Second Edition (WASI-II)*. Pearson: London, UK.

Zhang, S., & Malatesha Joshi, R. (2020). Longitudinal relations between verbal working memory and reading in students from diverse linguistic backgrounds. *Journal of Experimental Child Psychology*, *190*, 104727. https://doi.org/10.1016/j.jecp.2019.104727

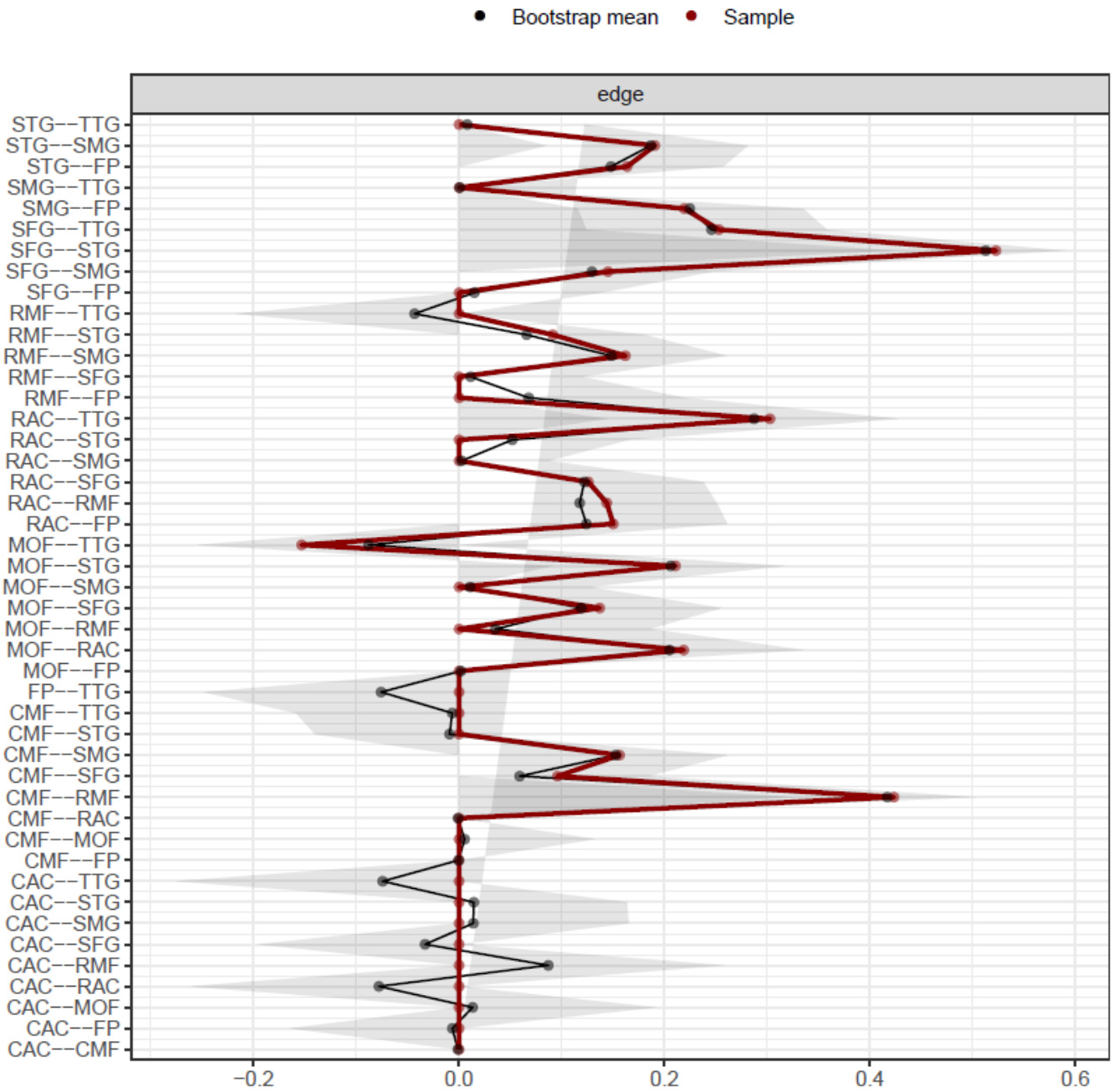
SUPPLEMENTARY MATERIAL

Edge-weight stability analyses

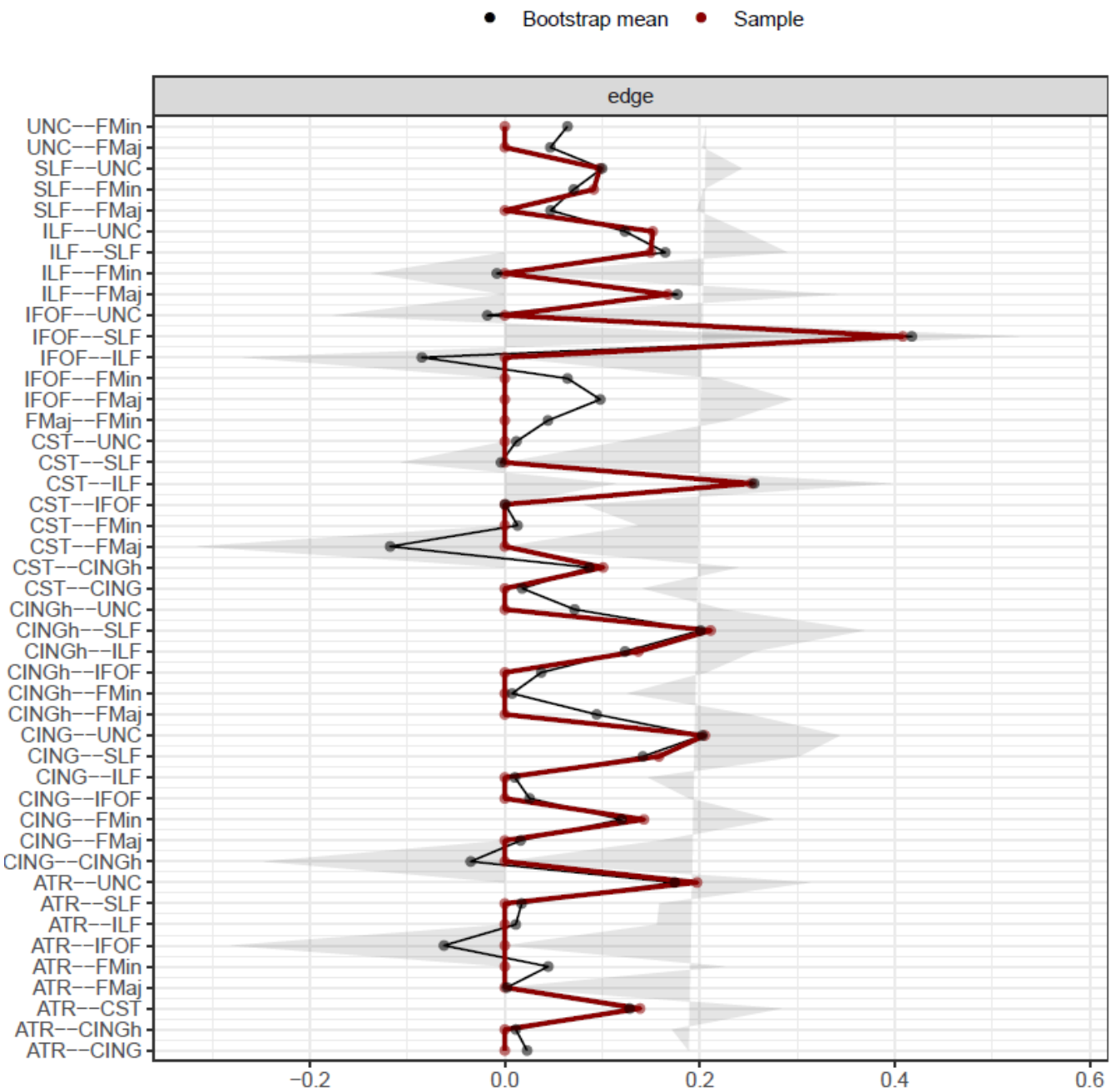
To further quantify the reliability of our partial correlation network edge-weights, we performed bootstraps (N=2,000) and compared the bootstrapped mean values to the original sample estimates (Supplementary Figures 1-3). We do not show the bootstraps for the multilayer networks due to the size of the plots but they (and all code for this project) can be found online (https://osf.io/36d2n/). Overall, most bootstrapped edge-weight means were consistently near the original sample value with the most variable being the tri-layer network (not shown), possibly due to the increase in complexity of the number of edge-weight parameters estimated.



Supplementary Figure 1. Comparisons between bootstrapped means and original sample edge-weight estimates for the CALM cognitive partial correlation network.



Supplementary Figure 2. Comparisons between bootstrapped means and original sample edge-weight estimates for the CALM grey matter partial correlation network.



Supplementary Figure 3. Comparisons between bootstrapped means and original sample edge-weight estimates for the CALM white matter partial correlation network.

How to deal with age?

As in previous literature, in our sample age shows a clear positive association with intelligence measures and brain structure (Figure 2). This fact, however, may further complicate any interpretations of (possible) causal interactions between cognitive and/or neural nodes. This leaves us with at least two options of how to deal with the relationship of age to cognitive ability, and grey and white matter structural covariance: 1) We could estimate the partial correlation network and include age as a node, therefore, choosing to estimate it *simultaneously* with the cognitive and neural variables (this is the option we chose for the main manuscript analyses), or 2). We could *regress out* the association of age for each variable (age would show no correlation for cognitive and/or neural measures) *before* network estimation. Both approaches have corresponding pros and cons. For instance, choosing to include age has the benefit of revealing the actual correlations among cognitive abilities and brain structure in the population. However, a drawback to this approach is that doing so could also amplify these associations, confounding our findings. On the other hand, regressing out age might enable us to detect correlations *beyond* age, possibility revealing core relations among variables independent of stereotypical neurocognitive development (e.g. older participants normally score higher on cognitive tasks and have larger brains as they mature). However, this might also remove developmental associations of interest (e.g. age may function as a moderator of cognitive and neural growth).

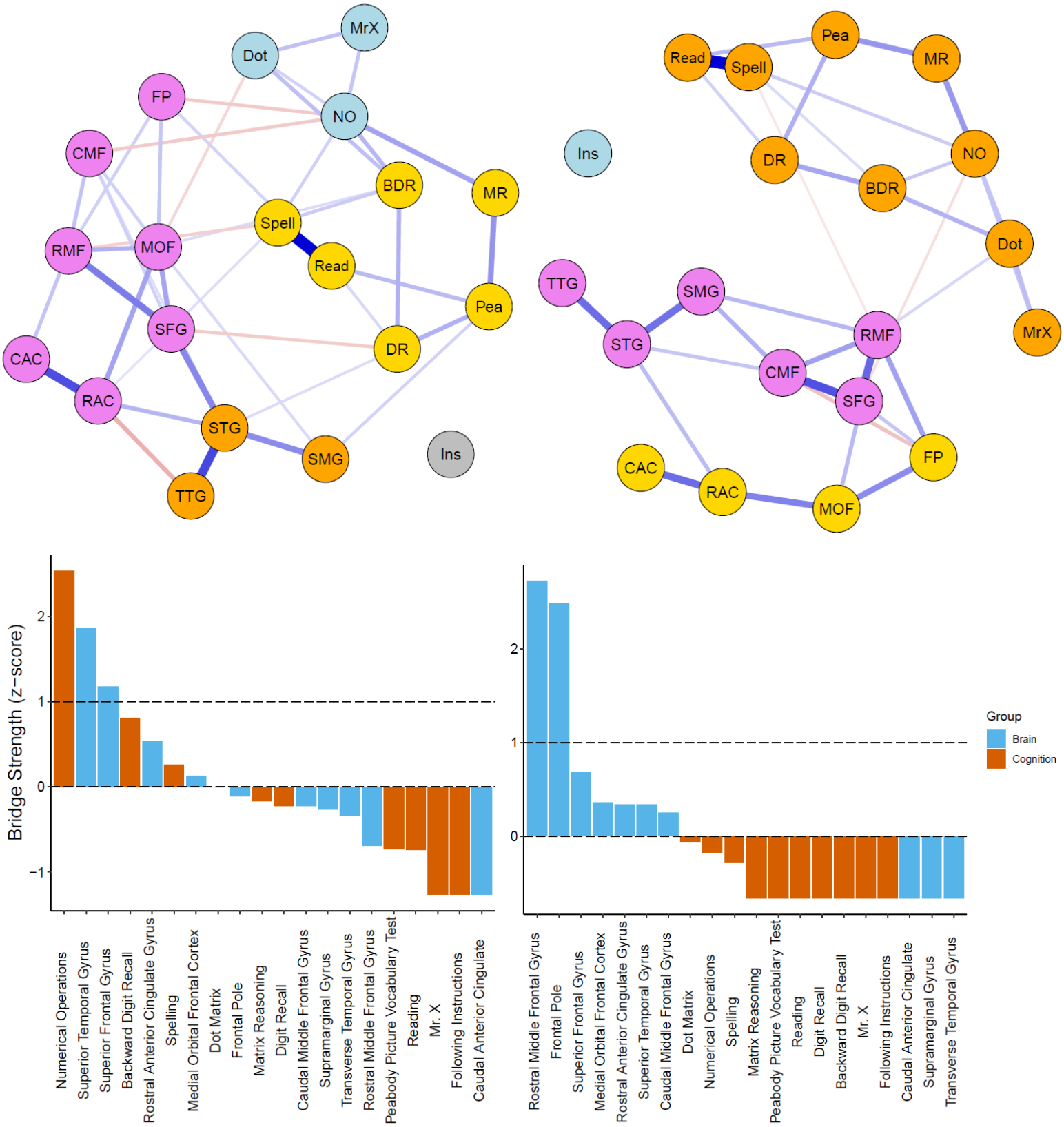
Here we compare the partial correlations matrices for the two analysis paths (age node used in network estimation vs. age node regressed out before estimation) for both single and multilayer networks (Supplementary Table 1). This analysis demonstrates that, regardless of how age is accounted for in estimation, the partial correlation networks are very similar to each other.

Network Type	Age Included in Estimation	Age Regressed Out before Estimation	Pearson Correlation
Cognitive	0.08 (0.11) [0, 0.63]	0.08 (0.12) [0, 0.65]	0.98
Grey Matter	0.09 (0.14) [-0.15, 0.52]	0.09 (0.14) [-0.15, 0.52]	1(rounded from 0.999)
White Matter	0.08 (0.11) [0, 0.44]	0.08 (0.13) [-0.2, 0.49]	0.93
Cognitive-grey matter	0.04 (0.1) [-0.12, 0.64]	0.03 (0.1) [-0.14, 0.65]	0.94
Cognitive-white matter	0.04 (0.1) [-0.2, 0.65]	0.03 (0.1) [-0.24, 0.66]	0.94
Tri-layer	0.02 (0.08) [-0.2, 0.66]	0.02 (0.07) [0, 0.64]	0.88

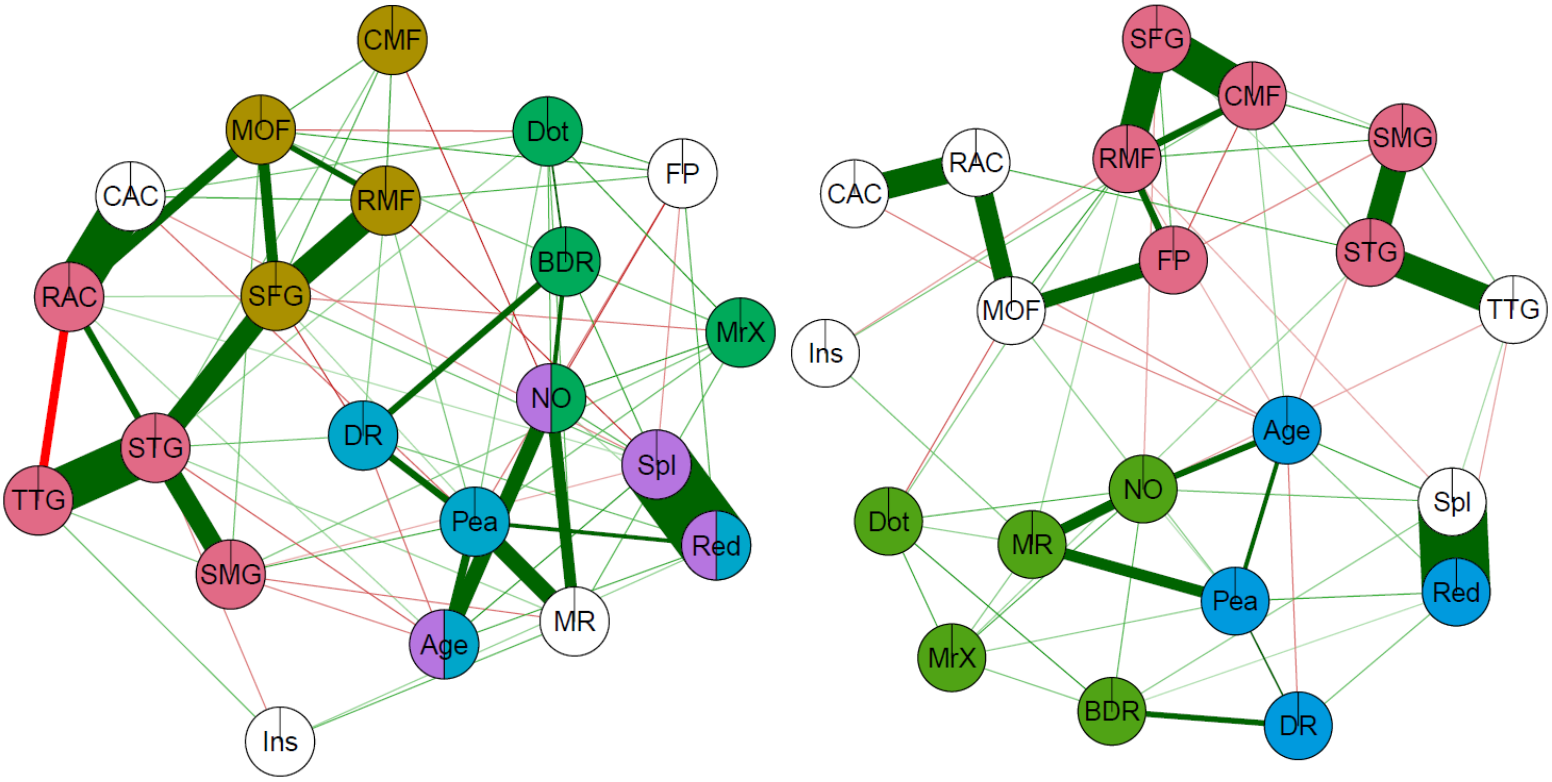
Supplementary Table 1. Comparisons between partial correlation networks (age included in estimation vs age regressed out before estimation). These include summary statistics such as mean, (standard deviation), [range] and Pearson correlations between PC graph weights using pairwise complete observations to account for missingness.

Teasing apart the relations of cortical volume to general intelligence: Multilayer analysis using cortical surface area and thickness

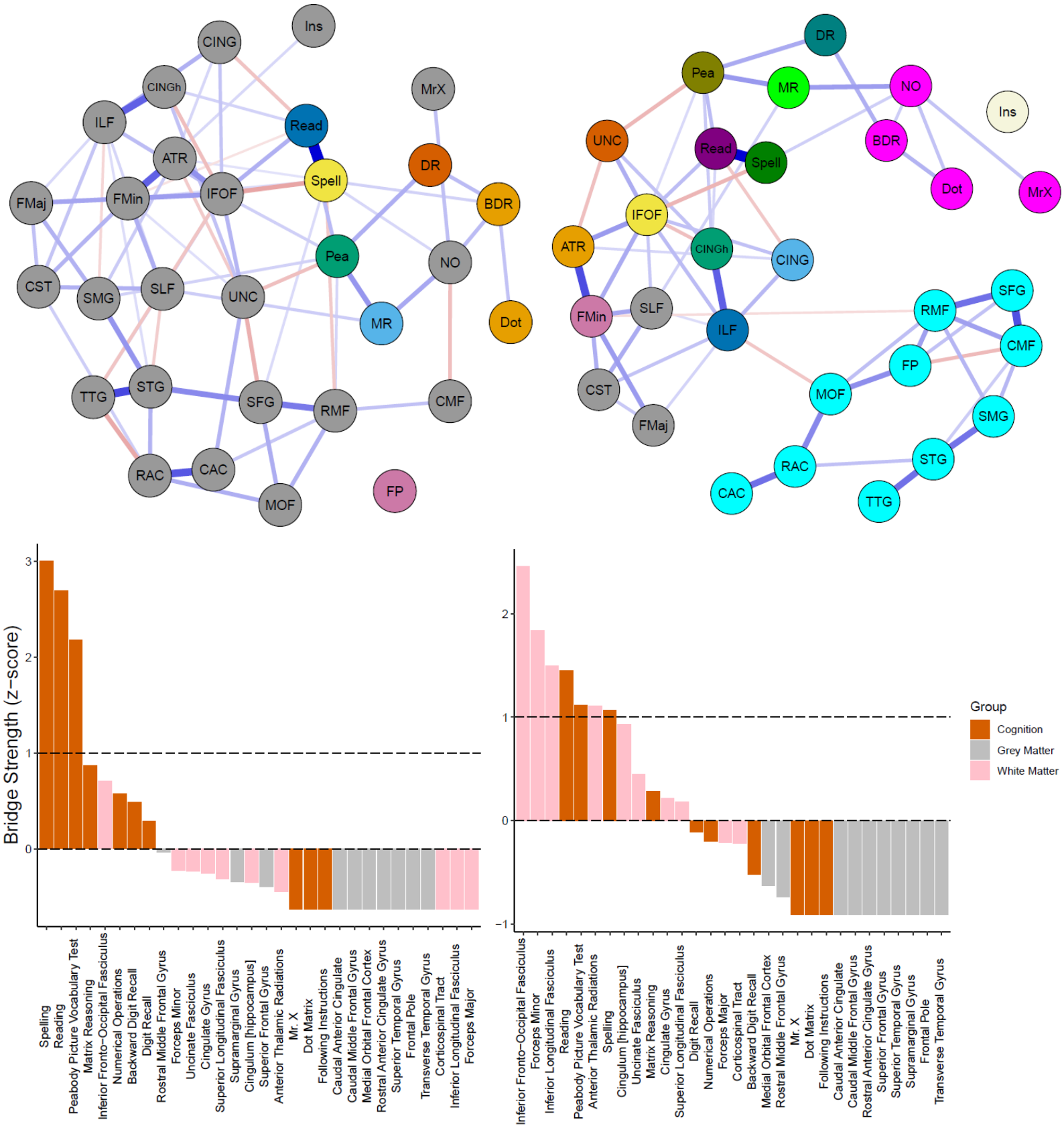
Lastly, we partitioned cortical volume into its constituent parts, cortical surface area and thickness, to compare their partial correlations and community structures when combined with white matter and general intelligence (Supplementary Figures 4-7). Although not conclusive, the effect seen for cortical volume in the main manuscript appears to be driven by cortical surface area, but not thickness, as exhibited by greater inter-connectivity among domains (brain vs behavior) and bridge strength distributed between both the brain and cognition (Supplementary Figure 4). Finally, when combined with fractional anisotropy, white matter appears to dominate the bridge strength centrality of the tri-layer network, especially for cortical thickness (Supplementary Figures 6 and 7).



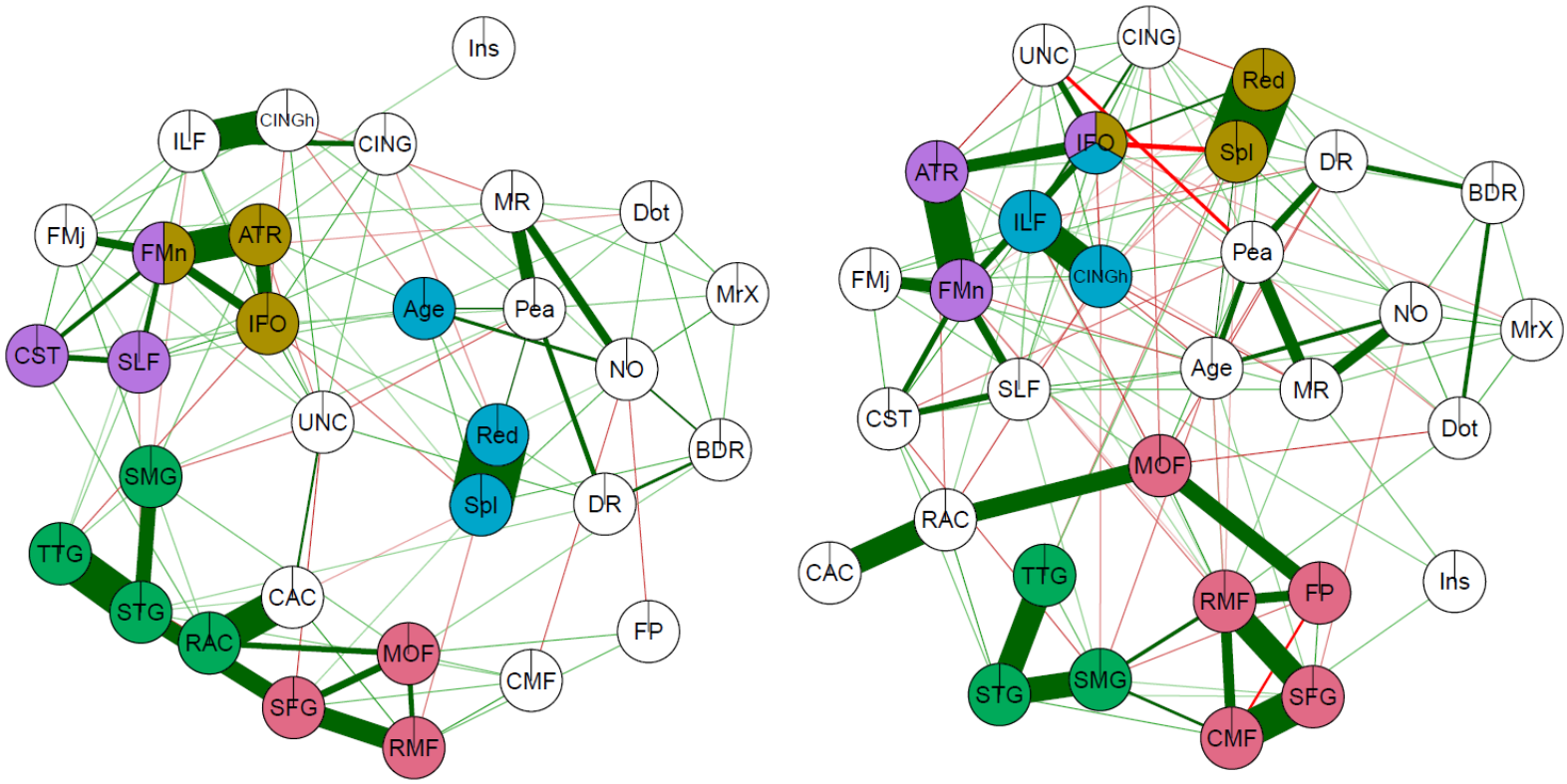
Supplementary Figure 4. Top: Network visualizations (spring layout) of partial correlation CALM bi-layer grey matter (surface area (left) and cortical thickness (right)) networks. Nodes are grouped according to Walktrap algorithm results (see above). Bottom: Bridge centrality estimates (z-scores) for CALM bi-layer grey matter (surface area (left) and cortical thickness (right)) networks. Dashed lines indicate mean strength and one standard deviation above the mean.



Supplementary Figure 5. Community detection (Clique Percolation) for cognition-surface area (left) and cognition-cortical thickness (right) networks. White-filled nodes indicate isolated nodes. Note that the 'Age' node is shown in these networks but not in the previous ones, although the estimation procedure was the same.



Supplementary Figure 6. Top: Network visualizations (spring layout) of partial correlation CALM tri-layer grey matter (surface area (left) and cortical thickness (right)) networks. Nodes are grouped according to Walktrap algorithm results (see above). Bottom: Bridge centrality estimates (z-scores) for CALM tri-layer grey matter (surface area (left) and cortical thickness (right)) networks. Dashed lines indicate mean strength and one standard deviation above the mean.



Supplementary Figure 7. Community detection (Clique Percolation) for tri-layer (surface area (left) and cortical thickness (right)) networks. White-filled nodes indicate isolated nodes. Note that the 'Age' node is shown in these networks but not in the previous ones, although the estimation procedure was the same.