Title: Modeling differences in neurodevelopmental maturity of the reading network using support vector regression on functional connectivity data

Running title: Reading disorder and network maturation

Authors: Oliver H.M. Lasnick[1*], Jie Luo[1], Brianna Kinnie[1], Shaan Kamal[2], Spencer Low[3], Natasza Marrouch[4], Fumiko Hoeft[1]

Affiliations:

[1]University of Connecticut, Storrs-Mansfield, CT, United States

[2]University of Connecticut School of Medicine, Farmington, CT, United States

[3]Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States

[4]University of Lisbon, Lisbon, Portugal

Acknowledgments: We thank the members of the Child Mind Institute for generously providing public access to their Healthy Brain Network Project dataset.

*Corresponding author (oliver.lasnick@gmail.com)

Abstract

The construction of brain growth charts trained to predict age based on morphometric and/or functional properties of the brain ('brain-age prediction') has the potential to inform neuroscientists and clinicians alike about children's neurodevelopmental trajectories. When applied to both typically- and atypically-developing populations – such as those with specific learning disorders – results may be informative as to whether a particular condition is associated with atypical maturation of specific brain networks. Here, we focus on the relationship between reading disorder (RD) and maturation of functional connectivity (FC) patterns in the prototypical reading/language network across development using a cross-sectional sample of N = 742 participants aged 6-21 years. A support vector regression model is trained to predict chronological age from FC data derived from (1) a whole-brain model, as well as (2) multiple 'reduced' models, which are trained on FC data generated from a successively smaller number of regions in the brain's reading network. We hypothesized that the trained models would show systematic underestimation of brain network maturity (i.e., lower FC-based age predictions) for poor readers, particularly for the models trained with reading/language regions. Exploratory results demonstrated that the most important whole-brain ROIs and connections are derived from the dorsal attention and somatosensory motor networks. Comparisons of the different models' predictions reveal that while the whole-brain model outperforms the others in terms of overall prediction accuracy, all models are effective at predicting brain maturity, including the one trained with the smallest amount of FC data. In addition, all models demonstrate some degree of moderation in the reliability of their age predictions as a function of reading ability, with predictions for both poor and exceptional readers being more accurate relative to those for typical readers.

Keywords: functional connectivity, fMRI, reading disorder, brain age prediction, support vector regression (SVR), network maturation

Key Points

- A machine learning model trained to predict age from functional connectivity data has better predictive utility for children/adolescents with either low OR high reading ability, and performs less well for those with average reading ability
- Both impaired and exceptional reading may be characterized as having lower variance in connectivity patterns on the population level compared to typical readers (rather than having delayed/accelerated network maturation)
- Neural markers of reading and language are likely diffusely represented in the brain

1. Introduction

1.1. Growth charts in neuroscience

For centuries pediatric growth charts – traditionally based on size metrics such as height, weight, and head circumference – have been invaluable in tracking childhood development (Cole, 2012). The use of analogous 'brain growth charts' allows neuroscientists to (1) estimate morphometric properties of the brain, such as regional gray/white matter (GM, WM) volumes and density in children, and (2) observe how these metrics change across time (Giorgio et al., 2008; Lemaitre et al., 2012; Bray et al., 2015). Using functional activity derived from magnetic resonance imaging (fMRI), which provides indirect measures of the brain's metabolic activity (Ogawa et al., 1993), further allows researchers to track age-related changes in neurophysiological processes. Examining both typical and atypical trajectories for brain growth and development is crucial for understanding the progression of neurodevelopmental disorders (Wang et al., 2021).

When it comes to growth charts derived from MRI data, researchers have previously tracked the maturational development of the default mode network (DMN) (Fair et al., 2008); and found differences in fMRI-based connectivity in attentional networks between children and adults (Fair et al., 2007). There have also been attempts to construct age-normed growth templates based on longitudinal structural data (Dong et al., 2020). Others have created machine learning (ML)-based growth curves from functional data: one of the first papers of this kind used an ML model known as a support vector machine, or SVM, to regress age on fMRI data [support vector regression (SVR)] (Dosenbach et al., 2010). Dosenbach et al. (2010)'s model was trained with functional connectivity (FC) data from resting-state fMRI (rs-fMRI) scans of both children and adults, with the resulting normative growth curve predicting neurological maturity. Based on this previous work, it is plausible that connectivity measures may be used to train an ML model to generate developmental growth

charts for a population with specific disorders – for instance, see Kessler et al. (2016), which focuses on brain network development in a population with attentional deficits. Such models may be useful in revealing the source of neurodevelopmental differences between typically- and atypically-developing populations. One area of research where this approach may be particularly useful is in the study of developmental dyslexia, a specific learning disorder defined by impairment in reading skills.

1.2. Neural correlates of dyslexia/RD and age-related development

Brain regions implicated in reading and language have been localized to an extensive and predominantly left-lateralized network in the brain. Reading network activity recruits left hemisphere ventral occipitotemporal, inferior frontal, and posterior parietal regions across both children and adults (Martin et al., 2015). Meta-analyses show that those with RD have reduced functional activity in regions overlapping with those commonly implicated during reading: inferior parietal cortex, the superior temporal gyrus, and the inferior frontal gyrus; reduced GM volume in the superior temporal gyrus, temporoparietal cortex, middle frontal gyrus, and superior occipital gyrus; and reduced WM in bilateral parieto-occipital regions (Maisog et al., 2008; Xia et al., 2016; Linkersdörfer et al., 2019). Bailey et al. (2018) also demonstrate that reading-related regions are highly distributed across the cortex, using the NeuroSynth meta-analytical database to derive regions most associated with reading-related activation (namely the temporal, inferior frontal, and supramarginal gyri, in addition to the temporo-parietal junction). Many of these dominant regions are also part of the frontoparietal attention network and the DMN. Furthermore, in Bailey et al. (2018) the modularity of visual, DMN, and cinguloopercular networks maintained the highest correlation with reading ability in subsequent analyses. In summary, there is a clear and widely-distributed neurobiological basis for RD(Peterson & Pennington, 2015; Swagerman et al., 2017).

There are also significant changes in functional activity that coincide with the transition from early childhood and adolescence to adulthood, and many of these involve regions implicated in the reading network. Functional resting-state studies have found that regions in the DMN shift from being sparsely interconnected during childhood to being densely interconnected by adulthood (Buckner et al., 2008; Fair et al., 2008). In the DMN, functional connections between the medial prefrontal cortex and both the posterior cingulate and lateral parietal regions have been shown to be stronger in adults compared to children. There are also major age-related changes in the dorsal and putative attentional networks (frontoparietal and cinguloopercular networks): in general, short-range connections tend to be negatively correlated with age (decrease in strength from childhood to adulthood), while long-range connections grow stronger with age (Fair et al., 2007). This latter phenomenon can be described as a developmental decrease in global modularity: an age-related decline in the degree of densely *intra*connected resting networks, accompanied by an increasing emphasis on long-range *inter*connected networks (Bailey et al., 2018). It is plausible that age-related changes in functional activity and connectivity may differ in children with and without reading deficits, and that these differences may be localized to regions involved in reading.

1.3. Applying ML-based growth charts to RD

To the best of our knowledge, there have been no studies to date that have specifically applied SVR methods to the construction of growth charts for children and adolescents with different levels of reading ability. A regression-based model that reliably predicts differences in brain development that are associated with reading-related skills could, in principle, be used to identify population-level deviation from developmental trajectories in children who might later go on to meet diagnostic criteria for dyslexia/RD. However, this possibility has yet to be fully explored in existing research.

For this study, the goal was to construct brain growth charts that distinguish the developmental trajectories of exceptional, typical, and impaired readers. RD has a prevalence of 5-10% in the general population and profoundly impacts one's ability to acquire advanced literacy skills, with cascading effects on academic achievement (Peterson & Pennington, 2015). In contrast, unexpectedly advanced decoding skills for one's age (often relative to comprehension skills and/or IQ) are often associated with autism spectrum disorders (ASD) or 'autism-like' traits (this profile is often referred to in the literature as 'hyperlexia'), while otherwise typically-developing children with advanced reading skills are referred to as 'precocious readers' (Ostrolenk et al., 2017). In this study, we include a sample of precocious readers without ASD as an additional control group, with the expectation that this sample might also show altered patterns of development in comparison to the typical-range and impaired readers (see Methods for additional details on each sample).

1.4. Aims and hypotheses

The primary aim of the current study is to examine, in a cross-sectional sample, whether there are differences between (a) children/adolescents with reading impairments, (b) precocious readers, and (c) typical readers in terms of estimated 'brain-age' based on FC data. We build on the prior literature by creating growth charts for brain networks in these reading subgroups, using ML models trained on FC data. Specifically, we (1) perform a preliminary literature search to isolate regions in the cortical reading network; and (2) use functional connectivity data from N = 742 children and adolescents aged 6-21 years (see section 2.2) to train an SVR model to predict chronological age for all groups, similar to Dosenbach et al. (2010). We repeat the procedure in (2) using a model trained on whole-brain connectivity data and 3 other 'reduced-ROI' models. Each reduced-ROI model is trained with a successively smaller feature set (generated from a smaller number of regions of interest [ROIs]) from the reading network derived in step (1). The goals are to: (i) compare age

predictions when using either distributed or specific reading network data; and (ii) determine if there is bias in the model based on reading ability, and if so, whether this effect is moderated by the cortical regions used to train the model. We hypothesize that the models will be biased in their age estimates for both the impaired readers and precocious readers compared to the controls, reflecting a developmental delay for the former and a developmental acceleration for the latter. Further, we expect that the whole-brain model will better predict age *in general* (collapsed across reading ability) compared to the reading-network trained models, due to being trained on more data, but that the reading-network models will emphasize functional differences between the groups. Therefore, group bias will be moderated by which data is used to train the model: reading group bias will be smaller in the whole-brain model compared to the reduced reading network-based models.

2. Materials and Methods

2.1. Methodological considerations

Resting-state data reflects the level of intrinsic activity within the brain while a participant is not engaged in cognitively demanding tasks. Regional correlations at rest are stable, generalizable, and may provide valuable data on differences in connectivity between controls and those with neuropsychiatric or neurodevelopmental disorders (Van Dijk et al., 2010). However, naturalistic audiovisual (AV) movie-watching paradigms may produce data with higher reliability than intrinsic resting-state and even task-based fMRI (Wang et al., 2017; Aliko et al., 2020). Moreover, the resting-state fMRI signal has been shown to be associated with increased functional connectivity within the DMN, a network implicated in typical development, aging, and reading (Fair et al., 2007). Therefore, there is both precedent and empirical support for the idea that resting-state data may reveal significant age-related functional networks that overlap with the developing reading network. Another important consideration for this study is that neuroimaging studies are often plagued by issues of low

sample size and selection bias, often due to strict exclusionary criteria (Turner et al., 2018; Kennedy et al., 2019; Specht, 2020). For these reasons, our analyses use resting-state functional connectivity data collected during a movie-watching paradigm (section 2.3), and we are less conservative in excluding participants who have comorbidities (section 2.2).

2.2. Demographics

2.2.1. Participants*.* All data is from the Child Mind Institute (CMI) Healthy Brain Network (HBN) Project, which collects neuroimaging and phenotypic data from a diverse sample of participants (Alexander et al., 2017). Neuroimaging data is publicly available for download on the CMI HBN website, while phenotypic data was made available for download via the Longitudinal Online Research and Imaging System (LORIS) upon completion of a Data Usage Agreement (DUA).

2.2.2. Inclusion criteria. We focused our analyses on participants who were at least 6 years of age at the time of data collection (age range 6-21 years). This age requirement was due to the use of scores from the Test of Word Reading Efficiency (TOWRE) for group classification criteria, as TOWRE is designed for ages 6-24 years (Torgesen et al., 2012). TOWRE involves two core subtests, in which the researcher presents a participant with a list of either real words or (pronounceable) pseudowords, which the participant is asked to read aloud one-by-one as quickly as possible within the allotted time. Real word reading is thought to index sight word reading ability, while pseudoword reading is thought to index phonemic decoding skills. A participant was identified as a poor reader (‘PR’) if they had a Total Word Reading Efficiency index score <=90, and as a typical reader (‘TR’) if they had a score >90 and <110 (91-109). Additionally, participants with a TOWRE index >=110 were determined to be precocious or ‘exceptional’ readers (‘ER’). Similar score cutoffs for TOWRE-based classification were identified in the sample from Graham et al. (2020). The TOWRE indices were calculated by those involved with the CMI HBN project using the TOWRE scoring manual. By-group

summed scores of the two TOWRE subtests [Phonemic Decoding Efficiency (PDE) and Sight Word Efficiency (SWE)] used to derive the TOWRE index for all participants are shown in Figure 1.

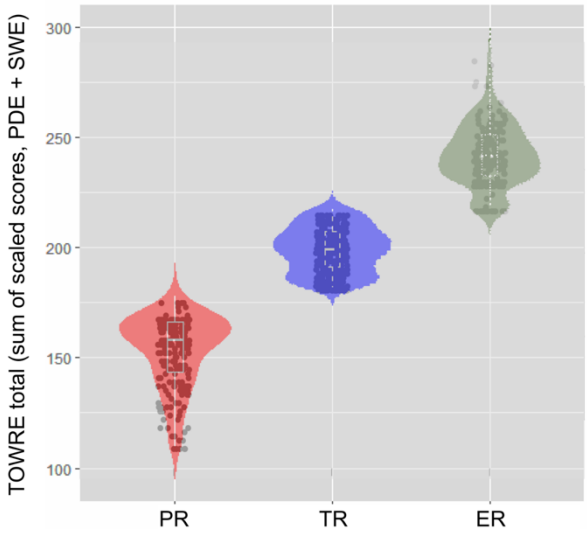


Figure 1. TOWRE total scores across each group.

2.2.3. Comorbidities. We took a liberal approach to previous neuropsychiatric and neurodevelopmental history in that all groups were allowed to have comorbid clinician diagnoses. The purpose of including comorbidities was to avoid the emergent population-level bias in traditional studies which disallow this kind of natural variation. It is well-known that RD has elevated rates of comorbidity with attention deficit hyperactivity disorder (ADHD) and other specific learning disorders, such as writing (dysgraphia) and mathematics (dyscalculia), as well as internalizing disorders such as depression and anxiety (Hendren et al., 2018). In addition, while RD has often been defined in terms of unexpectedly poor decoding in the absence of intellectual disability (Rutter & Yule, 1975), associations between RD and IQ scores (both verbal and nonverbal) have been reported at the population level, particularly for verbal IQ (van Bergen et al., 2014).

Table 1 contains all demographic and behavioral data on the participants after scan quality-control criteria had been applied (see 2.2). Those demographic variables with significant differences

between groups (as well as sex and handedness) were included as covariates (also called nuisance regressors) for all analyses, except for TOWRE subtest scores, which were used as grouping criteria; and full-scale IQ (FSIQ), which was highly correlated with performance (nonverbal) IQ (PIQ). The full set of covariates included in the statistical analyses were age, sex, handedness, PIQ, and SES.

Table 1. Demographic & behavioral data.

	ER N = 193	**TR N = 338**	**PR N = 211**	**Test statistic**	**Effect size (η^2)**
Age[1¶]	10.2 (2.9)$_a$	10.7 (2.7)$_{a,b}$	11.2 (3.6)$_b$	Welch's F(2,407.616) = 4.88 (*p* =.008)**	.015
Sex (% female)[¶]	29.5	36.7	34.6	X^2(2,739) = 2.81 (*p* =.245)	
Socioeconomic status (SES)[2¶]	9.0 (3.0)$_a$ [N=166]	8.5 (3.1)$_b$ [N=280]	6.8 (3.7)$_c$ [N=158]	Kruskal-Wallis H(2) = 37.85 (p < .001)***	.060
Handedness (% right)[¶]	91.7	86.7	91.0	X^2(2,739) = 4.19 (*p* =.123)	
FSIQ[1]	114.0 (14.4)$_a$	102.6 (12.1)$_b$	90.8 (12.6)$_c$ [N=178]	Welch's F(2,380.667) = 136.66 (*p* <.001)***	.298
PIQ (Block design, Scaled)[¶]	11.6 (3.1)$_a$	10.3 (3.2)$_b$	8.9 (2.9)$_c$ [N=178]	F(2,706) = 36.39 (*p* <.001)***	.093
TOWRE					
PDE (Scaled)[1]	118.0 (8.4)$_a$	98.4 (7.6)$_b$	76.1 (9.3)$_c$	Welch's F(2,411.073) = 1139.03 (*p* <.001)***	.778
SWE (Scaled)[1]	123.8 (8.6)$_a$	100.7 (7.8)$_b$	77.5 (10.0)$_c$	Welch's F(2,408.330) = 1269.40 (*p* <.001)***	.797

[1] Levene's test p < .05; Welch's statistic and Games-Howell correction used.
[2] Kruskal-Wallis nonparametric test reported for ordinal variable; Mann-Whitney U test for pairwise follow-ups.
¶Included as covariates for analyses.
$_{a,b,c}$ Cells with the same subscript indicate ns post-hoc group comparisons at the .05 alpha level.
Where data is missing, actual N is given in brackets.
All cells are in the format MEAN (SD) except for sex and handedness, which are given in percentages. IQ was measured using the Wechsler Intelligence Scale for Children – Fifth Edition (WISC-V). Post-hoc tests use Bonferroni correction for multiple comparisons if Levene's test is not significant (ns).

2.3. MRI data acquisition and processing

2.3.1. Resting fMRI. Data was collected during a movie-watching paradigm in which the participants viewed a 10-minute-long clip from the film *Despicable Me* while in-scanner. The 10-minute clip was presented as part of a longer in-scanner protocol and played at the following time stamp: 01:02:09–01:12:09 (HH:MM:SS). The protocol script was written in PsychoPy2 Experiment Builder v1.83.04 (Peirce, 2007).

2.3.2. Scan acquisition. Data (T1-weighted [T1w] structural scans and functional data) was collected at three possible scan locations. Acquisition parameters at each site were as follows:

(1) **Staten Island (SI).** 1.5T Siemens Avanto scanner (T1w: repetition time [RT] = 2730 ms, echo time [ET] = 1.64 ms, flip angle = 7°, 176 slices, 1.0×1.0×1.0mm voxels; fMRI: RT = 1450 ms, ET = 40 ms, flip angle = 55°, 54 slices, 2.5×2.5×2.5mm voxels).

(2) **Rutgers University Brain Imaging Center (RU)**. Siemens 3T Tim Trio scanner (T1w: RT = 2500 ms, ET = 3.15 ms, flip angle = 8°, 224 slices, 0.8×0.8×0.8mm voxels; fMRI: RT = 800 ms, ET = 30 ms, flip angle = 31°, 60 slices, 2.4×2.4×2.4mm voxels).

(3) **CitiGroup Cornell Brain Imaging Center (CG).** Siemens 3T Prisma scanner (T1w: RT = 2500 ms, ET = 3.15 ms, flip angle = 8°, 224 slices, 0.8×0.8×0.8mm voxels; fMRI: RT = 800 ms, ET = 30 ms, flip angle = 31°, 60 slices, 2.4×2.4×2.4mm voxels).

All data was downloaded from the CMI HBN Project neuroimaging data portal at http://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network/ (DOI: https://doi.org/10.1038/sdata.2017.181) and uses the brain imaging data structure (BIDS) format (https://bids.neuroimaging.io/). Data preprocessing and spatial registration of anatomical and functional data were performed using fMRIPrep v1.5.8 (Esteban et al., 2019). Additional denoising

and data extraction was performed in the CONN functional connectivity toolbox (Whitfield-Gabrieli & Nieto-Castanon, 2012).

The following information in sections 2.3.3-2.3.4 was copied from the boilerplate fMRIPrep output log after preprocessing was complete.

2.3.3. Anatomical data preprocessing (fMRIPrep). The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al., 2010), distributed with ANTs 2.2.0 (Avants et al., 2008), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), WM and GM was performed on the brain-extracted T1w using *fast* (FSL 5.0.9, Zhang et al., 2001). Volume-based spatial normalization to one standard space (MNI152NLin6Asym) was performed through nonlinear registration with antsRegistration (ANTs 2.2.0), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: ICBM 152 Nonlinear Asymmetrical template version 6 (TemplateFlow ID: MNI152NLin6Asym).

2.3.4. Functional data preprocessing (fMRIPrep). For each blood oxygen level dependent (BOLD) run, the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. A deformation field to correct for susceptibility distortions was estimated based on two echo-planar imaging (EPI) references with opposing phase-encoding directions, using 3dQwarp (AFNI 20160207). Based on the estimated susceptibility distortion, an unwarped BOLD reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using FLIRT (FSL 5.0.9) (Jenkinson & Smith, 2001) with the boundary-based registration

(Greve & Fischl, 2009) cost-function. Co-registration was configured with 9 degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using MCFLIRT (FSL 5.0.9) (Jenkinson et al., 2002). BOLD runs were slice-time corrected using 3dTshift from AFNI 20160207 (Cox & Hyde, 1997). The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as ‘preprocessed BOLD in original space’, or just ‘preprocessed BOLD’. The BOLD time-series were resampled into standard space, generating a preprocessed BOLD run in MNI152NLin6Asym space. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Several confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS, and three region-wise global signals. FD and DVARS are calculated for each functional run, both using their implementations in Nipype following the definitions by Power et al. (2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor, Behzadi et al., 2007). Principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 5% variable voxels within a mask covering the subcortical regions. This subcortical mask is obtained by heavily eroding the brain mask, which ensures it does not include cortical GM regions. For aCompCor, components are calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of

each functional run (using the inverse BOLD-to-T1w transformation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the *k* components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50% of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5mm FD or 1.5 standardized DVARS were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e., head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

2.3.5. Additional preprocessing in CONN. Regression of confounds from the BOLD signal was performed using the CONN functional connectivity toolbox (Whitfield-Gabrieli & Nieto-Castanon, 2012). At this stage, confounds (subcortical, WM, and CSF signals derived from tCompCor and aCompCor; framewise displacement; and realignment parameters) were regressed from the BOLD signal and high-motion frames were removed ('scrubbing', described in Power et al., 2012). A bandpass filter of 0.008-0.09 Hz was then applied to the signal, and linear detrending performed to correct for time-dependent artifactual signal drift.

2.3.6. Quality control & exclusion criteria. Participants were excluded from further analysis if (1) their collection of functional data did not include a full run of the movie-watching paradigm (for instance, if the participant left the scanner early and did not complete the full protocol), or (2) either their anatomical T1 scan or functional run was determined to be of insufficient quality. For T1 data, the MRIQC software's automatic classifier was used for exclusion. MRIQC's build-in classifier was run on each structural scan prior to fMRIPrep preprocessing; this algorithm automatically accepts or rejects each scan based on quality assessed using a vector of 64 features derived from 14 image quality metrics (Esteban et al., 2017). For functional data, the criteria for exclusion was having >⅓ of frames flagged for excess motion (>0.5mm FD) by fMRIPrep. This left a final sample size of N = 742 participants. The total percentage of ERs, TRs, and PRs excluded were 22.2%, 16.5%, and 26.2%, respectively. The percentage excluded due to functional motion outlier criteria was 3.6%, 6.7%, and 6.3%; the rest were due to insufficient structural data quality indicated by MRIQC. A Pearson Chi-squared test found that the rate of exclusion differed between the three groups: $X^2(2,939) = 9.77$ ($p = .008$). Post-hoc pairwise Chi-squared tests found that the only significant difference in exclusion was between the TRs and PRs: $X^2(1,691) = 9.62$ ($p = .002$), with the PRs having a higher rate of exclusion. The difference in exclusion rates between TRs and ERs ($X^2(1,653) = 3.21$; $p = .079$), and between PRs and ERs ($X^2(1,534) = 1.18$; $p = .312$), were ns.

2.4. ROI selection and feature construction for SVR training

2.4.1. ROI parcellation. This study utilized the Schaefer et al. (2018) local-global parcellation template to segment the brain for each participant into distinct ROIs (Fig 2a). The 400-ROI cortical parcellation, 7-network 2mm MNI version of the template was used. First, the template was loaded into the CONN toolbox along with functional data from all participants. For each participant, their BOLD data was partitioned into 400 ROI time series, with each time series being the average derived

from all voxels within a given ROI mask. These time series were used to calculate an ROI connectivity matrix for each participant: each value in the matrix is the Fisher-transformed bivariate correlation coefficient between a pair of ROI BOLD time series (Fig 2b). This generated a total of 400 *choose* 2 = $\frac{400!}{398!\ 2!}$= 79,800 possible unique pairwise features for each participant.

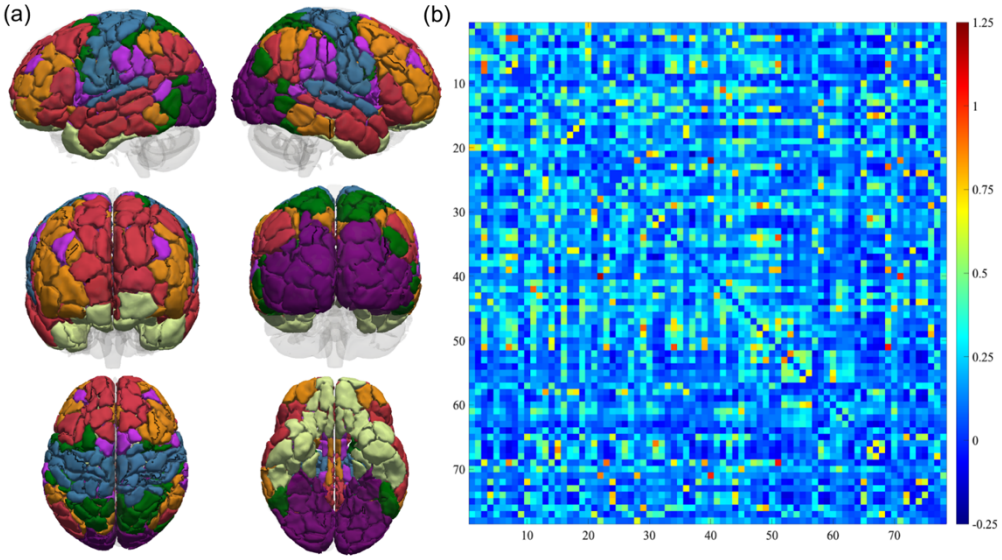


Figure 2. (a) Full 400-ROI, 7-network parcellation from Schaefer et al. (2018). (b) Example of an ROI-ROI connectivity matrix (reduced to 20% sparsity, 80 ROIs).

2.4.2. ROI selection. We next performed a literature search for meta-analyses on neuroimaging studies related to reading and language abilities in children and adults. The aim was to generate a set of coordinates in MNI space which could then be mapped to ROIs in the Schaefer parcellation. Papers were included if they met the following criteria: (1) examined studies performed with structural MRI/fMRI or PET imaging; (2) focused on group contrasts between participants with and without a reading/language impairment, or a functional contrast between reading/language and control tasks; (3) reported results as coordinates in either Talairach or Montreal Neurological Institute (MNI) space (describing relevant foci for each group/task); and (4) focused on English-speaking participants or cross-cultural results. Talairach coordinates were converted to MNI using the

Lancaster transformation implemented in GingerALE (Laird et al., 2010; Eickhoff et al., 2009; Eickhoff et al., 2012; GingerALE software: https://www.brainmap.org/ale/). A total of 27 meta-analyses were examined. Of these, 22/27 reported at least a subset of their significant coordinates. (Four papers did not report significant coordinates. One paper was determined to be unrelated due to its focus on semantic and syntactic processing in typical readers.)

From the included 22 papers, 416 unique coordinates were identified as being part of the reading/language network based on group or functional task contrasts. Each coordinate was then mapped to the ROI in the Schaefer template in which that coordinate was spatially contained (referred to as a given coordinate's 'parent ROI' in the Schaefer template). 49/416 were out of the template's range (mapped onto subcortical regions or the cerebellum), leaving a total of 367 in-range coordinates. These 367 coordinates loaded onto a total of 152 unique Schaefer ROIs (some coordinates mapped onto the same ROI; see Fig 3). The Supplementary Material includes a spreadsheet with all information used for the literature search, including full citations for the included meta-analyses and their corresponding coordinates.

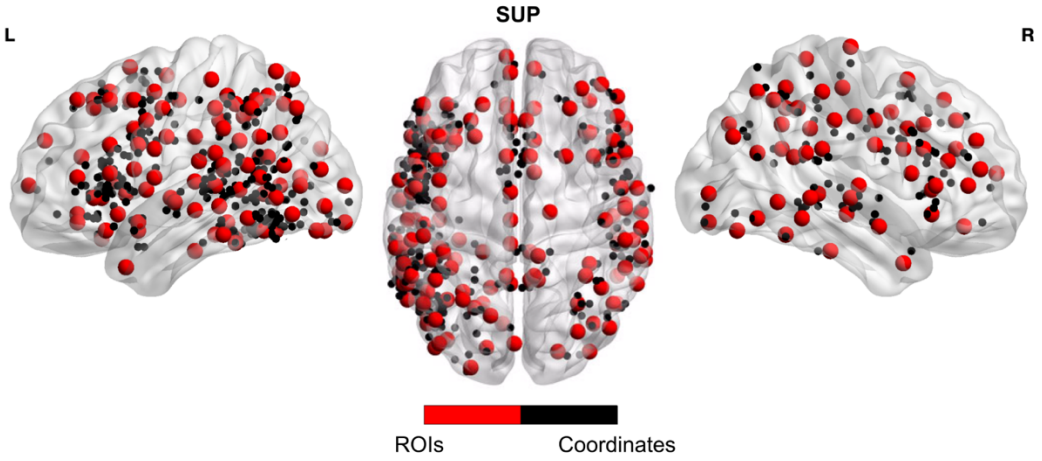


Figure 3. ROIs identified by the literature review. L = Left view; SUP = Superior; R = Right. Black dots represent the 367 coordinates reported across 22 meta-analyses (1mm sphere). Red dots

represent the centroid coordinates for the corresponding 152 ROIs in the Schaefer parcellation; ROI centroids are drawn with a 5mm-radius sphere. A given coordinate's 'parent ROI' is defined as the ROI in which that coordinate is encompassed.

2.4.3. ROI cutoffs. To investigate the specific contribution of reading and language networks to brain-age predictions, 4 ROI cutoffs were chosen, each with a successively smaller number of ROIs. Cutoffs were based on the total number of coordinates that mapped to a given Schaefer ROI. We argue that the greater the number of unique coordinates from the meta-analyses that map to a given region, the more likely that region is to be part of the core language/reading network and thus relevant to group differences within our participant populations. Using this criterion, we created the following sets of ROIs: 400 ROIs (whole-brain/no cutoff, Fig 4a); 152 ROIs (≥1 coordinates which have that region as its parent ROI, Fig 4b); 78 ROIs (≥2 coordinates, Fig 4c); and 27 ROIs (≥4 coordinates, Fig 4d). For the subsequent analyses, pairwise ROI-ROI connections were derived only for the included ROIs in each cutoff. As the ROI cutoff increases, the data used to train the models thus becomes more concentrated in regions of the brain often implicated in reading and reading deficits: in particular, left hemisphere temporal, parietotemporal, occipitotemporal, inferior frontal, and precentral regions and homologous right hemisphere regions, although the left hemisphere becomes more dominant as the cutoff increases (Fig 4, c-d).

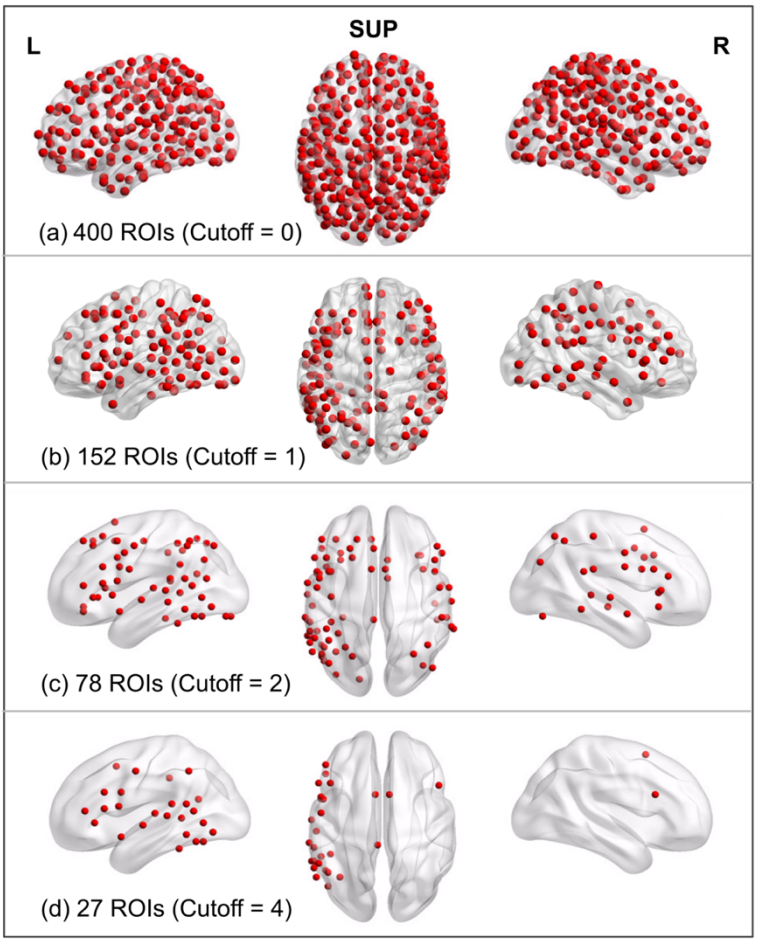


Figure 4. ROI sets generated from literature review. L = Left view; SUP = Superior; R = Right. 4 sets of ROIs ranging from (a) whole-brain 400-ROI set, to (d) reduced 27-ROI reading network set. Centroids are drawn with a 5mm-radius sphere.

2.4.4. Principal component analysis (PCA). Prior to SVR model training, PCA was applied to all functional connectivity data in the training set to reduce the number of features used in training while preserving data variance. PCA is a common dimensionality reduction technique that reduces the number of dimensions in a dataset by applying a linear orthogonal transformation. This converts the feature set into a smaller number of principal components (PCs), where each component is a linear combination of several correlated original features (Leonardi et al., 2013; Amico & Goñi, 2018). The set of features for a dataset is thus projected onto a lower-dimensional space, and all final components are uncorrelated with one another. Not all components need be retained in the final

transformed feature set, as the derived linear transformation is computed from the top eigenvectors of the covariance matrix and provides a ranking of the components in terms of which explain the most variability in the data (Khosla et al., 2019). A specified amount of variance can therefore be retained by discarding components which do not contribute much additional information. For each set of ROIs (whole-brain and reduced) PCA was performed using the built-in functionality in Python's Scikit-learn toolbox and supplied with an argument to retain 95% of the variance in the original data.

2.4.5. Model training, cross-validation, & testing. Each SVR model was built using Python's Scikit-learn toolbox. A linear kernel was chosen, and hyperparameter ε was set to 0.1 such that the model received no penalty in the training loss function for data points predicted within 0.1 years of the true age value. Pre-PCA features were all pairwise ROI-ROI Fisher's Z-transformed bivariate correlation coefficients between the included regions for that model's set of ROIs. The final set of training features consisted only of the principal components which remained after running PCA (Fig 5). Training labels were floating-point numbers describing chronological age for all participants.

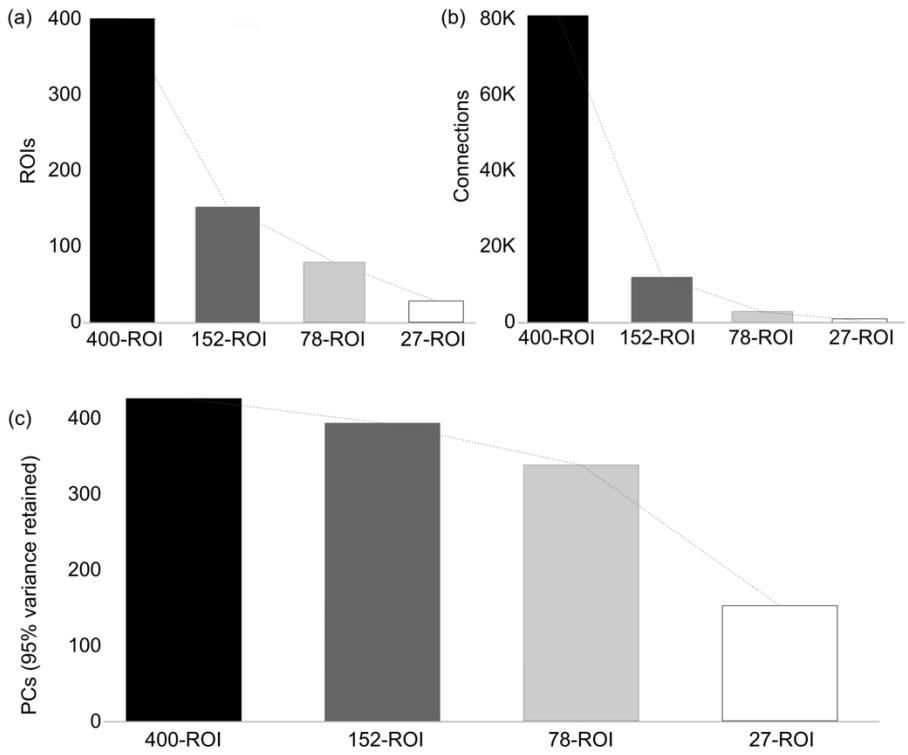


Figure 5. (a) The number of ROIs in each cutoff model; (b) The number of initial pairwise functional connections for each model prior to PCA = N choose 2 = $\frac{N!}{(N-2)!\ 2!}$, where N is the number of ROIs; (c) The number of principal components used to train each model after preserving 95% of data variance from b).

For each model, 1000 training permutations were done, each with a unique split for training and testing data. For each permutation data was split into the training and test sets pseudorandomly using a 2-1 train-test ratio (496 training participants, 246 testing participants). The pseudorandom split was performed for all three groups (PR, TR, ER) separately to ensure that each group was represented proportionally in both the training and testing sets. PCA was performed during each permutation on the training data only, to ensure that the test set was withheld from all models throughout training and cross-validation. The resulting PCA model was used to transform the training data prior to training/cross-validation, and the same model was used to transform the test set prior to testing. Within each permutation 5-fold cross-validation was done on the training set.

After training and testing was complete, each model had generated 246 x 1000 [(# testing samples per permutation) x (# permutations)] = 246,000 total predictions. It is worth noting that, although the training and testing sets were generated via a pseudorandom split for each permutation, the chance of a given participant not being in the testing set at least once is extremely low; indeed, the number of times each participant was placed in the test set is approximately uniform.

2.5. Analyses

2.5.1. Demographic covariates. Prior to running the planned analyses, we tested for group differences in the following demographic variables: age, sex, SES, handedness, and WISC-V FSIQ and PIQ (Block design [BD] subtest). We tested continuous variables using one-way ANOVA with Bonferroni correction for multiple comparisons. If Levene's test was significant below the .05 alpha level, we reported Welch's F-statistic and used Games-Howell correction for multiple comparisons. SES, an ordinal variable coding for annual income brackets, was compared using the nonparametric Kruskal-Wallis test and Mann-Whitney U tests as pairwise follow-ups. We tested categorical variables (sex and handedness) using Chi-squared tests. See Table 1 in section 2.1 for a summary of results.

We found a small-to-medium group difference in FSIQ (η^2 = .298), a smaller effect for PIQ (η^2 = .093), and a small effect for SES (η^2 = .060), using effect size interpretations outlined in Sawilowsky (2009). Previous findings have shown that (a) both verbal and nonverbal measures of IQ are associated with and predictive of later reading ability, including dyslexia diagnoses (van Bergen et al., 2014); that (b) verbal IQ measures are more strongly related to reading and language deficits compared to nonverbal measures (van Bergen et al., 2014); and that (c) RD is disproportionately more likely to be identified in children from low-SES environments (Piotrowska & Willis, 2019). Our findings here are directly supportive of (a)-(c). This provides additional support for previous demographic findings in the RD population and justifies these variables' inclusion as nuisance

450 regressors in subsequent analyses. We selected PIQ instead of FSIQ as the IQ regressor in the
451 analysis because FSIQ is more likely to be associated with reading/language skills, and using it as a
452 regressor/covariate might eliminate important variation between reading groups. Since we anticipate
453 differences in these skills between good and poor readers and wish to preserve these trends in our
454 data, we therefore selected PIQ. Neither handedness nor biological sex differed between groups and
455 were also included as covariates in all analyses.

456 There was also a very small effect of age (η^2 = .015). Post-hoc tests showed that the
457 significant difference in age was driven by the age gap between ERs and PRs (p <.05), with poor
458 readers being older on average, while the TRs were intermediate in age and not different from the
459 other two in post-hoc follow-ups (both ps >.05).

460 **2.5.2. Construction of brain growth charts and identification of the most important ROIs**
461 **/ connections [Whole-brain SVR, exploratory].** First, an SVR model is trained on PCs derived from
462 functional connectivity data associated with the whole-brain 400-ROI parcellation, as described in
463 2.3.5. After training and testing is complete, there are 1000 permutations' worth of prediction results.
464 For each of these permutations, the PCA coefficient matrix (N x M, where N = number of PCs, and M
465 = number of original ROI-ROI connections) and SVR model coefficient array (1 x N) are matrix-
466 multiplied to get an average coefficient for each original ROI-ROI connection (1 x M). These 1 x M
467 arrays are then averaged across all permutations to get a final average coefficient for each ROI-ROI
468 connection. The coefficients were ranked and sorted by absolute value, and the top 10% (7,980
469 connections) were selected for further examination. The distribution of all feature coefficients from the
470 whole-brain model is approximately normal (Fig 6). As each coefficient is the linear combination of
471 that ROI-ROI connection's weights in the PCA matrix and the SVR model's training weight for each
472 principal component, its absolute value is taken to represent that feature's relative contribution to

overall model predictions. In addition, the top 10% of features were decomposed into their component ROIs. The component ROIs were then pooled together to determine which occurred most frequently in the highest-ranked features.

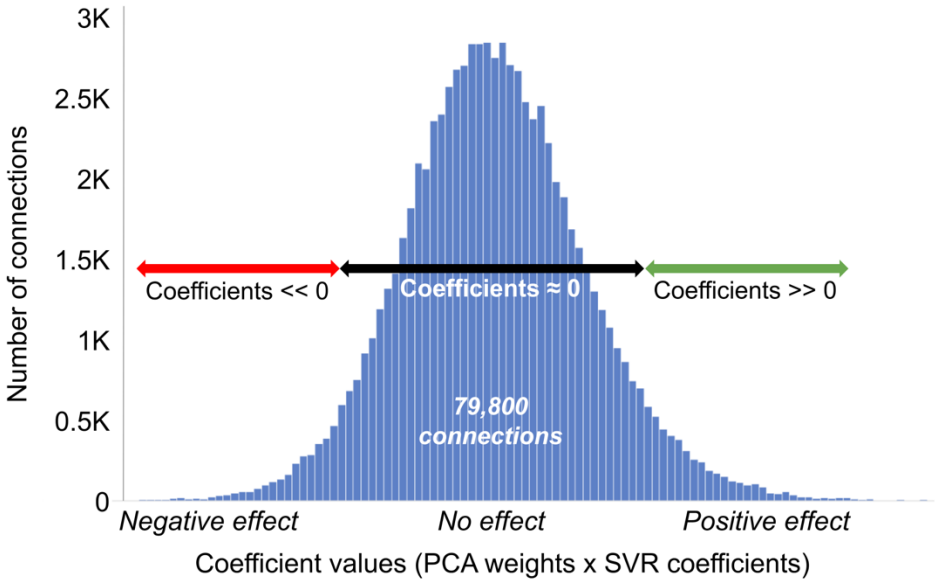


Figure 6. Histogram of coefficients for ROI-ROI connections in the whole-brain (400-ROI) model.

2.5.3. Between-model differences in age prediction [Whole-brain SVR vs. reduced-ROI SVR]. To compare prediction accuracy between the whole-brain and reduced-ROI reading network models, we performed the following steps. Predictions were averaged for each participant across all permutations, for a final set of 742 predictions. All models were then compared on their ability to correctly predict chronological age *across all participants,* regardless of decoding ability, using Akaike's Information Criteria (AIC). AIC estimates a model's prediction error. It can be used to directly compare the predictive ability of non-nested linear statistical models by calculating their relative likelihood. We hypothesized that there would be a positive relationship between model performance (minimized prediction error = smaller AIC) and the number of ROI connections (prior to PCA) used to train the model. Therefore, models trained with more (whole-brain) data would more accurately predict the chronological age of the participants, collapsed across all three reading groups.

2.5.4. Between-group differences in SVR predictions at different ROI cluster levels. To test for group bias in each model, we ran a series of linear regression analyses with True Age (t-Age) and Group as the independent variables, and Predicted Age (pr-Age) as the dependent variable. Sex, handedness, SES, and PIQ were included as nuisance regressors. An interaction between t-Age and Group (slope bias) indicates that the model's predictive ability varies by reading phenotype. A main effect of Group (intercept bias) suggests that the model systematically under- or overestimates age based on reading phenotype. We expected that the trained models would underestimate PRs' age relative to the TRs and ERs, suggesting a relative developmental delay in functional connectivity in the poor readers. We therefore tested whether there was an effect of Group or a t-Age x Group interaction on model predictions for each ROI cutoff. We expected that as more ROIs are excluded from model training, there would be a larger intercept bias, attributable to systematic underestimation of PRs' ages in models that have been trained with more reading and language ROIs. Specifically, we expected that there would be a larger difference in pr-Age accuracy based on the participants' reading group (main effect of Group), but that overall model fit would be largely preserved (small or no effect of t-Age x Group).

3. Results

3.1. Whole-brain SVR

ROIs. The whole-brain model was trained to predict age using PCs derived from 79,800 ROI-ROI connections generated from 400 ROIs. Section 2.4.2 describes how each connection's relative weight (magnitude and direction) in the overall SVR model was derived. The individual weights for all 79,800 connections are shown in Figure 7 below.

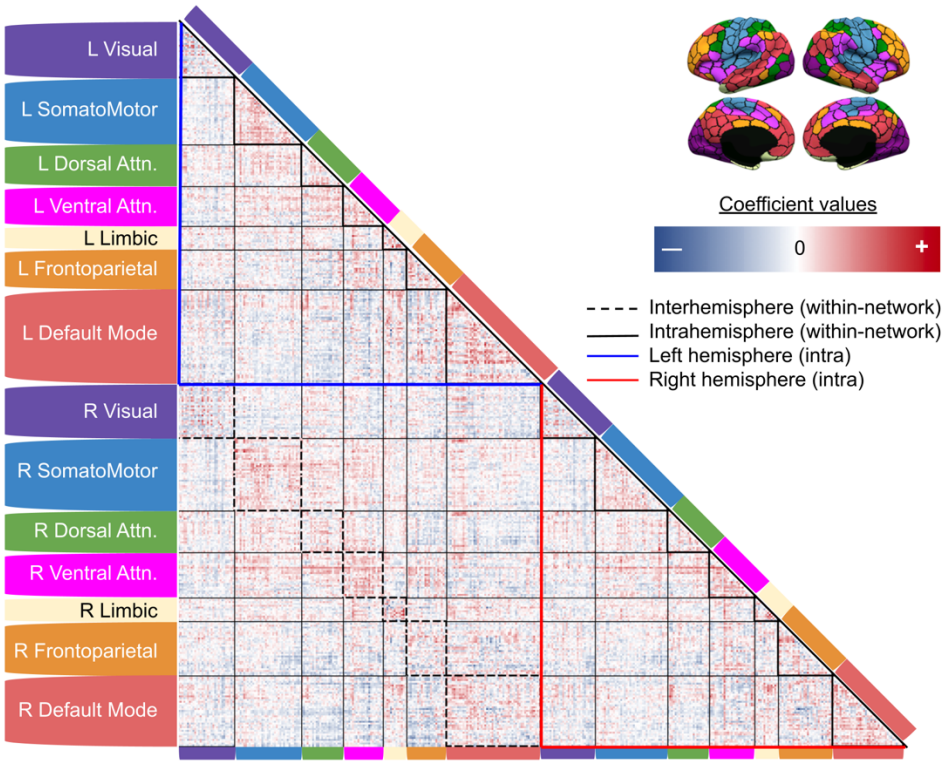


Figure 7. Model weights for all pairwise connections in whole-brain model (SVR x PCA). Each box or triangle enclosed by black lines reflects the connectivity between two brain networks (sets of ROIs), using the Yeo et al. (2011) 7-network parcellation as described in Schaefer et al. (2018). Thicker black solid lines (along the diagonal) indicate intrahemispheric and intra-network connections. Thicker dotted lines are interhemispheric, intra-network connections, reflecting the connectivity between the homologous networks of each hemisphere. The blue triangle (upper left quadrant) reflects connectivity between all ROIs in the left hemisphere; the red triangle (lower right quadrant) reflects connectivity between all ROIs in the right hemisphere.

For the individual ROI/connection analysis, the top 10% of connections were decomposed into their component ROIs. These component ROIs were then ranked in terms of frequency of occurrence across the top connections. Across all 400 ROIs, the median rate of occurrence in the top 10% connections was 39 [range 5-74, standard deviation (SD) = 8.99]. The total number of ROIs with a frequency greater than one SD above the median (≥ 48) was N = 76 ROIs. Of these, 41/76 (53.9%) were from the right hemisphere. When the top ROIs were restricted to a frequency greater than two SDs above the median (≥ 57), 13/16 (81.3%) of these ROIs were from the right hemisphere;

particularly from the right dorsal attention network (three ROIs) and the somatosensory motor network (four ROIs). Two others were from the right temporal pole; another two were from the right visual cortex. The final two right-lateralized ROIs came from the default mode network and the parietal lobe. Overall, the highest ranked ROIs were more likely to be right-lateralized, with the exceptions of a left-medial ROI from visual cortex and two more strongly left-lateralized ROIs from the somatosensory network and temporal pole, respectively. Overall, the highest-ranked ROIs came from the right dorsal attention network, bilateral visual areas, right somatosensory motor areas, and bilateral temporal poles (Fig 8).

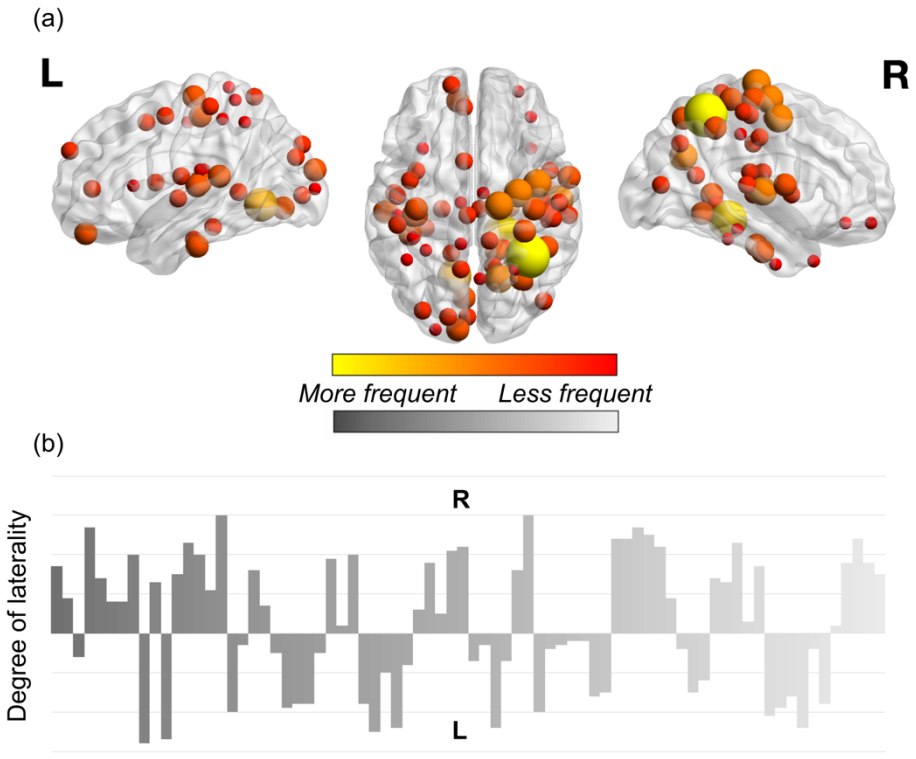


Figure 8. Top ROIs in the whole-brain model. ROIs are from the top 10% of connections and occur with a frequency >1 SD above the median. (A) Location of top ROIs in an MNI template brain. (B) Degree of hemispheric lateralization for each ROI, from most frequent (darker) to least frequent (lighter). Bar height is equal to the ROI's x-coordinate in MNI space (medial ≈ 0). L = Left, R = Right.

Connections. A disproportionate amount of the highest-weighted connections were interhemispheric (long-range) rather than intrahemispheric (local). This suggests that the strength of

connectivity between long-distance regions is a more salient feature associated with age and/or developmental changes in functional connectivity, as opposed to connection strength between short-range local ROIs.

Network analysis. To further examine the relationship between age and connectivity between brain regions, the whole-brain model's weighting of network interconnectivity was additionally explored. Each of the 400 ROIs defined in Schaefer et al. (2018) is assigned to one of seven networks (with both left and right homologs) described in the Yeo at al. (2011) parcellation. For these networks, the average SVR weights for between-network connectivity were calculated as follows: for two networks "A" and "B", the set of model coefficients for all ROI-ROI features (in which one ROI is from "A" and the other from "B") is averaged. Additionally, averages of positive model coefficients and negative model coefficients were calculated separately to determine which inter-network connections had the most extreme weights (Fig 9).

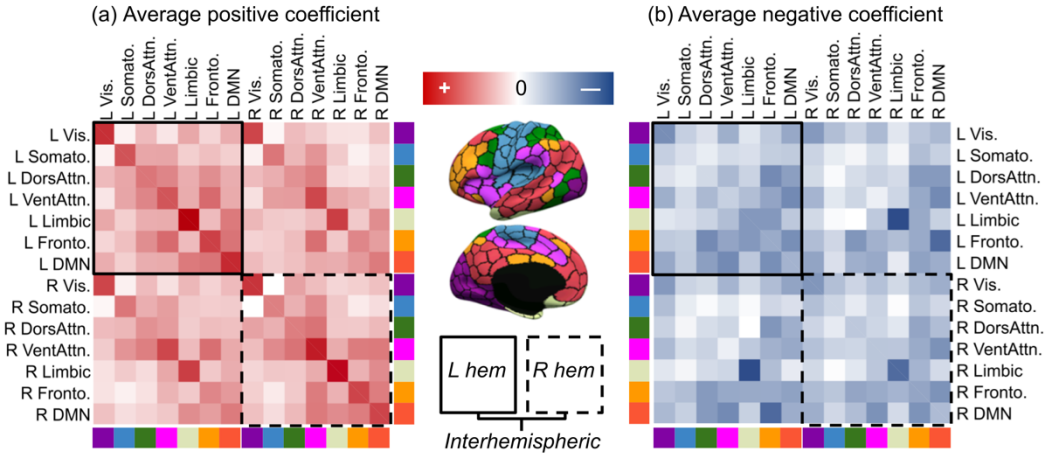


Figure 9. Network analysis for whole-brain model. Each matrix value reflects an average coefficient value for all the pairwise ROI-ROI connections between two networks. (a) Average positive coefficients. (b) Average negative coefficients. Networks are defined using the Yeo et al. (2011) 7-network parcellation described in Schaefer et al. (2018).

Results here indicate that connectivity *within* networks – both in the same hemisphere and between left and right network homologs tend to have more extreme weights (Fig 9a and 9b). The

connections with the strongest weights are: within-network connections of the left and right limbic networks; within-network connections of the left and right visual networks; within-network connections of the left frontoparietal network; within-network connections of the left DMN; interhemispheric connections between the left and right limbic network; long-range connections between the right DMN and left frontoparietal network; and within-network connections of the right limbic network.

3.2. Between-model differences in age prediction [Whole-brain SVR vs. reduced-ROI SVR]

Best-fit regression lines for all models and corresponding R^2 are shown in Figure 10. AIC results showed that the whole-brain model minimized prediction error for chronological age (Table 2). Thus, we can be confident that the whole-brain model captures a meaningfully larger amount of age-related variance in the functional data.

Table 2. Prediction accuracy for all models (whole-brain and reduced-ROI).

df	Model	$R^2_{tot,\ adj}$	$t_{pr\text{-}Age}$	$\beta_{pr\text{-}Age}$	p	$SE_{pr\text{-}Age}$	AIC_i	$RL_{i,\ min}$
F(1,740)	400-ROI	.343	19.67	.586	<.001	.032	3452.082	1
	152-ROI	.271	16.62	.521	<.001	.035	3528.733	2.27×10^{-17}
	78-ROI	.197	13.53	.445	<.001	.039	3600.215	6.81×10^{-33}
	27-ROI	.154	11.66	.394	<.001	.040	3639.077	2.48×10^{-41}

All regression models with a single predictor, predicted age (pr-Age), for the dependent variable, true age (t-Age). $\beta_{pr\text{-}Age}$ are standardized coefficient values. $RL_{i,\ min}$ values are relative likelihoods of model *i* compared to the model which minimized AIC. SE = Standard error. df = degrees of freedom.

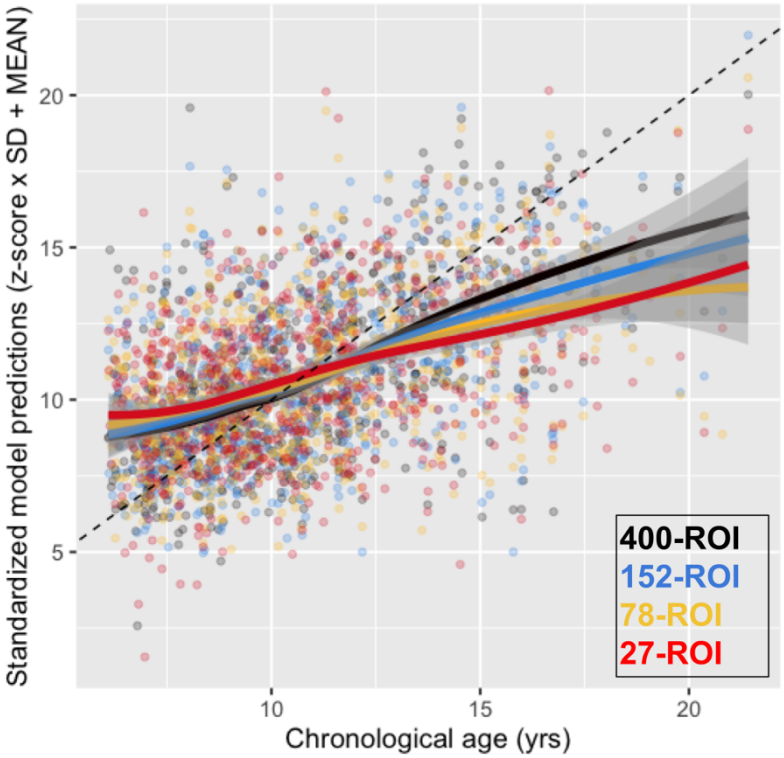


Figure 10. Visualization of relationship between true age (x-axis) and predicted age for each ROI model (y-axis). The dotted reference line is at $y = x$, reflecting hypothetical "perfect" age prediction.

3.3. Between-group differences in SVR predictions at different ROI cluster levels

Statistical results of the regression analyses for each ROI cutoff are shown in Table 3. True age (t-Age) accounted for a significant proportion of variance in predicted age (pr-Age) across all four reduced-ROI models when controlling for covariates (see effects of t-Age and corresponding R^2_{tot}).

Table 3. Effects on SVR predictions across reduced-ROI models.

	df	R^2_{tot}	t	β	ΔR^2	*p*
400 ROIs (whole-brain)						
t-Age	5,582	.305	**15.17**	.529	.305	***<.001***
Group (PR vs. TR)	7,580	.306	0.50	.019	.001	.615
Group (ER vs. TR)			-0.41	-.016		.682
Group x t-Age (PR vs. TR)	9,578	.330	**2.09**	.086	.024	***.037***
Group x t-Age (ER vs. TR)			**4.53**	.203		***<.001***
152 ROIs						

t-Age	5,582	.251	**12.82**	.464	.251	***<.001***
Group (PR vs. TR)	7,580	.251	-0.02	-.001	.000	.986
Group (ER vs. TR)			-0.46	-.018		.645
Group x t-Age (PR vs. TR)	9,578	.275	**2.11**	.090	.024	***.035***
Group x t-Age (ER vs. TR)			**4.24**	.198		***<.001***
78 ROIs						
t-Age	5,582	.184	**10.57**	.400	.184	***<.001***
Group (PR vs. TR)	7,580	.187	-0.86	-.036	.003	.393
Group (ER vs. TR)			-1.57	-.064		.118
Group x t-Age (PR vs. TR)	9,578	.209	1.85	.083	.022	.064
Group x t-Age (ER vs. TR)			**4.01**	.195		***<.001***
27 ROIs						
t-Age	5,582	.134	**8.67**	.337	.134	***<.001***
Group (PR vs. TR)	7,580	.136	-0.17	-.007	.002	.868
Group (ER vs. TR)			-0.86	-.037		.388
Group x t-Age (PR vs. TR)	9,578	.148	1.13	-.005	.012	.260
Group x t-Age (ER vs. TR)			**2.86**	-.026		***.004***

R^2_{tot} includes all effects of interest plus covariates (sex, SES, handedness, PIQ). Effects for t-Age control only for covariates, while Group and Group x t-Age effects additionally control for t-Age. The model including the Group x t-Age effect additionally controls for Group. ΔR^2 values are equal to the change in total R^2 values when that predictor is added to the model. df = degrees of freedom.

Contrary to our hypothesis, the main effect of Group on model predictions was non-significant across all ROI clusters, lowest in the whole-brain model, and highest in the 78-ROI model. At the same time, the interaction effects of Group x Age were significant across all ROI clusters, being highest in the whole-brain model and decreasing successively along with the number of ROIs used for initial feature construction (Fig 11).

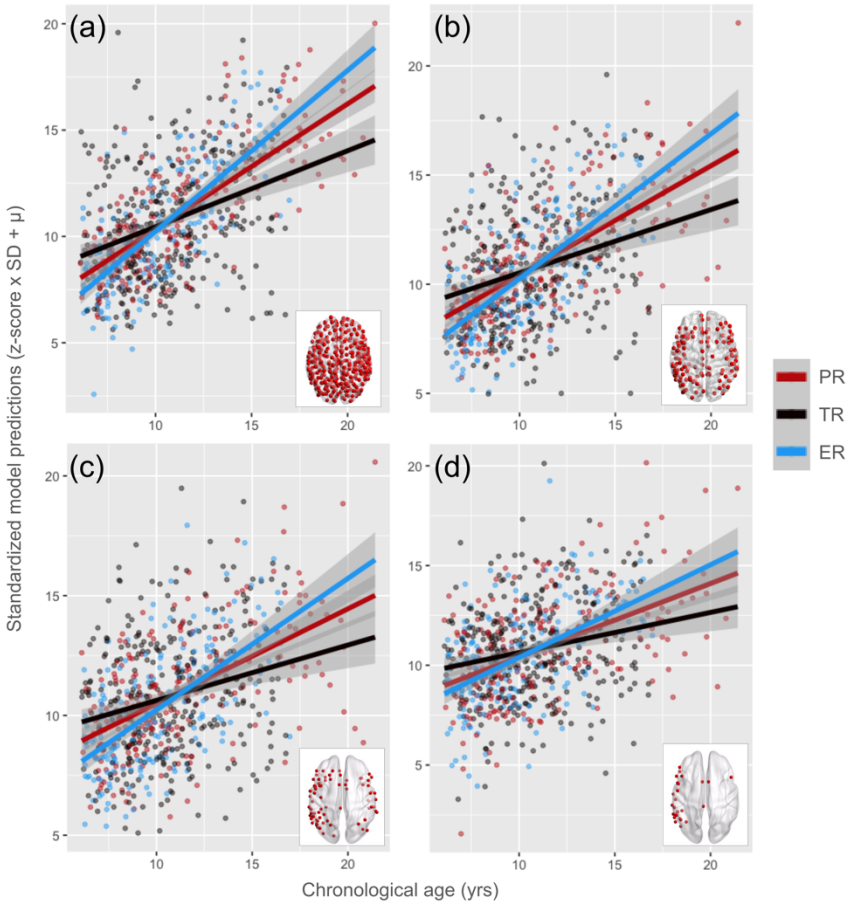


Figure 11. Relationship between true age and model age estimates based on FC data for each ROI cutoff. (a) 400 ROIs. (b) 152 ROIs. (c) 78 ROIs. (d) 27 ROIs. ROIs used to train each model are shown in the bottom right corner of each subplot. Line/dot color indicates reading group.

4. Discussion

In summary, we first performed a preliminary analysis to determine which functional connections and individual brain regions were the most heavily weighted in the whole-brain SVR model; we then performed our main analyses, for which we predicted that (1) the whole-brain model would be a better predictor of age, on average, than the reduced-ROI models, and (2) there would be a group bias in model age predictions when trained on functional connectivity from reading network ROIs (reduced-ROI models). The preliminary analyses showed that for the whole-brain model, (i) connections involving right-hemisphere ROIs from the dorsal attention and somatosensory motor

networks were the most heavily weighted in predicting age, and that (ii) in general, within-network connections (both intra- and interhemispheric) were more highly weighted than between-network ones. The main analyses confirmed the hypothesis that the whole-brain model would be a better predictor of age overall than the reduced-ROI models; however, there was no apparent prediction bias due to reading status in any of the models. On the contrary, there were differences in model fit due to reading status, with both poor and exceptional readers having more reliable age predictions than typical readers; this effect was largest in the whole-brain model.

4.1. Exploratory whole-brain SVR

The exploratory results suggest that while the maturation of functional networks is widely distributed across the brain, regions and connections with the most predictive value come from the right dorsal attention network, bilateral sensory regions, and the temporal lobes. The heaviest-weighted connections are longer-range interhemispheric connections, rather than short-range within-network intrahemispheric ones. Broadly, interhemispheric prefrontal connections have a marked negative effect on predicted age, while interhemispheric connections between temporal and occipital regions have a positive effect. Longer-range intrahemispheric connections - particularly between regions within the frontal lobe and between prefrontal and parietal regions - also tend to grow stronger with age.

4.2. Between-model differences in age prediction

For the main analyses, our most straightforward expectation was that the whole-brain model would be a better overall predictor of age than the other three models, which was indeed supported by our results. Its predictions accounted for far more variance in the participants' true age and it had a much higher relative likelihood than the other models, as indicated by the AIC results.

633 It is worth noting that the level of variance explained between the whole-brain model collapsed
634 across groups (R^2 = .343) and the most reduced-ROI model with 27 ROIs (R^2 = .155) decreased only
635 by approximately half, despite the models' features being constructed from vastly different amounts of
636 FC data. The total number of connections used to construct features for the 27-ROI model (351
637 pairwise connections) was <.005% of the amount used for the whole-brain model (79,800
638 connections). Yet, this model was still able to capture a significant amount of variance in age based
639 on connectivity between a highly restricted set of brain regions. This suggests that while training with
640 widely-diffused whole-brain data will result in more accurate estimates for age, the degree of
641 accuracy is not linearly proportional to the total number of ROIs or connections used to generate
642 features. Markers of age, in addition to markers of reading ability, are present and detectable in
643 connectivity patterns across the brain.

644 **4.3. Between-group differences in SVR predictions**

645 We also predicted that there would be a larger group bias in our model's age predictions when
646 it was trained to predict age using FC data from reading network ROIs rather than diffuse whole-brain
647 connectivity data. The ROIs on which the reduced-model connections were based were drawn
648 preferentially from left-hemisphere reading and language regions (see Fig 4, b-d). Contrary to
649 predictions, there was no apparent bias solely due to group status in any of the ROI cluster models.
650 When the features used to train the SVR model came predominantly from reading and language
651 ROIs, differences in model performance between the groups did not become more pronounced.

652 However, the significant Group x Age interactions, which were strongest in the whole-brain
653 model, introduce a potential alternative indicator of developmental abnormalities in functional
654 connectivity for readers of both high and low reading ability. Rather than supporting the presence of a
655 simple group bias, these results suggested that the reliability (model slope, or R^2) of age predictions

656 varies as a function of reading ability in all models. In other words, model fit was better for the ERs
657 (R^2 = .566) and PRs (R^2 = .558) compared to the TRs (R^2 = .118) (see individual linear slopes in Fig
658 11). This effect was strongest in the whole-brain model, which was trained with more features derived
659 from diffuse regions of cortex. This suggests that characteristic connectivity patterns associated with
660 reading ability may be diffused widely across the brain rather than localized. One interpretation of this
661 result is that the variability of FC patterns is reduced in both ERs and PRs relative to TRs. This may
662 hinder the ability of our SVR model to reliably predict age within the TR population to the same extent
663 as within the other groups. A possible follow-up to these results may be to examine whether the same
664 functional connections identified as driving the model's predictions in our exploratory analysis also
665 differ between these reading groups. Such an analysis could utilize a more traditional GLM-based
666 statistical approach, using both resting-state and task-based data.

667 **4.4. Limitations and Future Directions**

668 There are several limitations to the current study, both conceptual and methodological. The
669 most obvious methodological drawback is that the SVR model coefficients are agnostic towards the
670 initial sign of the correlation between two ROIs' activity. Bearing this in mind, we do not make any
671 strong claims about the precise developmental trajectories of highly-weighted connections, but have
672 rather focused our exploratory analysis on the specific regions and networks that are most heavily
673 weighted in the trained model regardless of their initial sign or the signs of their associated SVR
674 coefficients. Future studies which aim to examine these developmental trajectories (and their
675 differences in controls vs. those with RD) more closely may consider using either a traditional group-
676 based analytical approach or a simple linear regression model trained on a smaller number of
677 features from the most developmentally-salient networks, with the latter decided *a priori* based on
678 previous literature.

Another significant limitation is conceptual. An obvious point of criticism is that we have trained the SVR model to predict age using a representative split from the PR, TR, and ER populations. Therefore, the model was trained to predict the age of poor (and exceptional) readers nearly as much as it was trained to predict the age of controls. This approach may have 'washed out' meaningful differences between these groups that might have emerged had the model only been trained with control participants. An alternative approach that uses the latter method of training only with controls would address this limitation.

Furthermore, we do not assess the model's ability to predict reading skills directly. A follow-up study that instead assesses the model's ability to predict continuous TOWRE (or other reading assessment) scores rather than age would be more enlightening about which functional networks in the brain underlie reading ability. This approach may also address the issue of age differences between phenotypic groups since there would no longer be any group designations based on reading scores; regardless, age would likely still need to be included as a control variable, since we have noted that chronological age is associated even with age-standardized TOWRE scores in our sample. An alternative approach would be to split participants into narrower age bands after training, and then examine whether there are group differences in the model's age estimates at these different age bins. This would allow us to not only compare age-matched subgroups but would also be informative as to whether there are any developmental patterns to these differences (e.g., present in early childhood but then diminish later in life).

5. Conclusion

In line with previous literature, this study has illustrated that models trained on functional connectivity data are able to reliably predict the true age of participants. Contrary to our expectations, our results do not support the hypothesis that individuals with RD/dyslexia have delayed

reading/language brain network maturity compared to controls or advanced readers. Our analysis suggested that both poor and exceptional readers exhibit more reliable age predictions compared to controls, implying that these groups may be characterized by less variability in their connectivity patterns. Future research should focus on refining these predictive models and predicting reading skills as opposed to age, which may be informative as to which ‘age-agnostic’ brain networks underlie reading development. The integration of machine learning with neuroimaging, as demonstrated in this study, provides promising avenues for advancing diagnostic and intervention strategies for children with neurodevelopmental and learning disorders of all kinds.

References

Alexander, L. M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., Vega-Potler, N., Langer, N., Alexander, A., Kovacs, M., Litke, S., O'Hagan, B., Andersen, J., Bronstein, B., Bui, A., Bushey, M., Butler, H., Castagna, V., Camacho, N., Chan, E., … Milham, M. P. (2017). An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Scientific data, 4*, 170181. https://doi.org/10.1038/sdata.2017.181(Giorgio et al., 2008)

Aliko, S., Huang, J., Gheorghiu, F., Meliss, S., & Skipper, J. I. (2020). A naturalistic neuroimaging database for understanding the brain using ecological stimuli. *Scientific data, 7*(1), 347. https://doi.org/10.1038/s41597-020-00680-2

Amico, E., & Goñi, J. (2018). The quest for identifiability in human functional connectomes. *Scientific reports, 8*(1), 8254. https://doi.org/10.1038/s41598-018-25089-1

Bailey, S. K., Aboud, K. S., Nguyen, T. Q., & Cutting, L. E. (2018). Applying a network framework to the neurobiology of reading and dyslexia. *Journal of neurodevelopmental disorders, 10*(1), 37. https://doi.org/10.1186/s11689-018-9251-z

Bray, S., Krongold, M., Cooper, C., & Lebel, C. (2015). Synergistic Effects of Age on Patterns of White and Gray Matter Volume across Childhood and Adolescence. *eNeuro, 2*(4), ENEURO.0003-15.2015. https://doi.org/10.1523/ENEURO.0003-15.2015

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences, 1124*, 1–38. https://doi.org/10.1196/annals.1440.011

Cole T. J. (2012). The development of growth references and growth charts. *Annals of human biology, 39*(5), 382–394. https://doi.org/10.3109/03014460.2012.694475

Dosenbach, N. U., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., Nelson, S. M., Wig, G. S., Vogel, A. C., Lessov-Schlaggar, C. N., Barnes, K. A., Dubis, J. W., Feczko, E., Coalson, R. S., Pruett, J. R., Jr, Barch, D. M., Petersen, S. E., & Schlaggar, B. L. (2010). Prediction of individual brain maturity using fMRI. *Science (New York, N.Y.), 329*(5997), 1358–1361. https://doi.org/10.1126/science.1194144

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage, 59*(3), 2349–2361. https://doi.org/10.1016/j.neuroimage.2011.09.017

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. H*uman brain mapping, 30*(9), 2907–2926. https://doi.org/10.1002/hbm.20718

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PloS one, 12*(9), e0184661. https://doi.org/10.1371/journal.pone.0184661

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature methods, 16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Fair, D. A., Cohen, A. L., Dosenbach, N. U., Church, J. A., Miezin, F. M., Barch, D. M., Raichle, M. E., Petersen, S. E., & Schlaggar, B. L. (2008). The maturing architecture of the brain's default network. *Proceedings of the National Academy of Sciences of the United States of America, 105*(10), 4028–4032. https://doi.org/10.1073/pnas.0800376105

Fair, D. A., Dosenbach, N. U., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., Barch, D. M., Raichle, M. E., Petersen, S. E., & Schlaggar, B. L. (2007). Development of distinct control networks through segregation and integration. *Proceedings of the National Academy of Sciences of the United States of America, 104*(33), 13507–13512. https://doi.org/10.1073/pnas.0705843104

Giorgio, A., Watkins, K. E., Douaud, G., James, A. C., James, S., De Stefano, N., Matthews, P. M., Smith, S. M., & Johansen-Berg, H. (2008). Changes in white matter microstructure during adolescence. *NeuroImage, 39*(1), 52–61. https://doi.org/10.1016/j.neuroimage.2007.07.043

Hendren, R. L., Haft, S. L., Black, J. M., White, N. C., & Hoeft, F. (2018). Recognizing Psychiatric Comorbidity With Reading Disorders. *Frontiers in psychiatry, 9*, 101. https://doi.org/10.3389/fpsyt.2018.00101

Kennedy, D. N., Abraham, S. A., Bates, J. F., Crowley, A., Ghosh, S., Gillespie, T., Goncalves, M., Grethe, J. S., Halchenko, Y. O., Hanke, M., Haselgrove, C., Hodge, S. M., Jarecka, D., Kaczmarzyk, J., Keator, D. B., Meyer, K., Martone, M. E., Padhy, S., Poline, J. B., Preuss, N., … Travers, M. (2019). Everything Matters: The ReproNim Perspective on Reproducible Neuroimaging. *Frontiers in neuroinformatics, 13*, 1. https://doi.org/10.3389/fninf.2019.00001

Kessler, D., Angstadt, M., & Sripada, C. (2016). Growth Charting of Brain Connectivity Networks and the Identification of Attention Impairment in Youth. *JAMA psychiatry, 73*(5), 481–489. https://doi.org/10.1001/jamapsychiatry.2016.0088

Khosla, M., Jamison, K., Ngo, G. H., Kuceyeski, A., & Sabuncu, M. R. (2019). Machine learning in resting-state fMRI analysis. *Magnetic resonance imaging, 64*, 101–121. https://doi.org/10.1016/j.mri.2019.05.031

Laird, A. R., Robinson, J. L., McMillan, K. M., Tordesillas-Gutiérrez, D., Moran, S. T., Gonzales, S. M., Ray, K. L., Franklin, C., Glahn, D. C., Fox, P. T., & Lancaster, J. L. (2010). Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: validation of the Lancaster transform. *NeuroImage, 51*(2), 677–683. https://doi.org/10.1016/j.neuroimage.2010.02.048

Lemaitre, H., Goldman, A. L., Sambataro, F., Verchinski, B. A., Meyer-Lindenberg, A., Weinberger, D. R., & Mattay, V. S. (2012). Normal age-related brain morphometric changes: nonuniformity across cortical thickness, surface area and gray matter volume?. *Neurobiology of aging, 33*(3), 617.e1–617.e6179. https://doi.org/10.1016/j.neurobiolaging.2010.07.013

Leonardi, N., Richiardi, J., Gschwind, M., Simioni, S., Annoni, J. M., Schluep, M., Vuilleumier, P., & Van De Ville, D. (2013). Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest. *NeuroImage, 83*, 937–950. https://doi.org/10.1016/j.neuroimage.2013.07.019

Linkersdörfer, J., Jurcoane, A., Lindberg, S., Kaiser, J., Hasselhorn, M., Fiebach, C. J., & Lonnemann, J. (2015). The association between gray matter volume and reading proficiency: a longitudinal study of beginning readers. *Journal of cognitive neuroscience, 27*(2), 308–318. https://doi.org/10.1162/jocn_a_00710

Maisog, J. M., Einbinder, E. R., Flowers, D. L., Turkeltaub, P. E., & Eden, G. F. (2008). A meta-analysis of functional neuroimaging studies of dyslexia. *Annals of the New York Academy of Sciences, 1145*, 237–259. https://doi.org/10.1196/annals.1416.024

Martin, A., Schurz, M., Kronbichler, M., & Richlan, F. (2015). Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Human brain mapping, 36*(5), 1963–1981. https://doi.org/10.1002/hbm.22749

Ogawa, S., Menon, R. S., Tank, D. W., Kim, S. G., Merkle, H., Ellermann, J. M., & Ugurbil, K. (1993). Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophysical journal, 64*(3), 803–812. https://doi.org/10.1016/S0006-3495(93)81441-3

Peirce J. W. (2007). PsychoPy--Psychophysics software in Python. *Journal of neuroscience methods, 162*(1-2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce J. W. (2009). Generating Stimuli for Neuroscience Using PsychoPy. *Frontiers in neuroinformatics, 2*, 10. https://doi.org/10.3389/neuro.11.010.2008

Peterson, R. L., & Pennington, B. F. (2015). Developmental dyslexia. *Annual review of clinical psychology, 11*, 283–307. https://doi.org/10.1146/annurev-clinpsy-032814-112842

Piotrowska, B., & Willis, A. (2019). Beyond the global motion deficit hypothesis of developmental dyslexia: A cross-sectional study of visual, cognitive, and socio-economic factors influencing reading ability in children. *Vision research, 159*, 48–60. https://doi.org/10.1016/j.visres.2019.03.007

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage, 59*(3), 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018

Richlan F. (2020). The Functional Neuroanatomy of Developmental Dyslexia Across Languages and Writing Systems. *Frontiers in psychology*, *11*, 155. https://doi.org/10.3389/fpsyg.2020.00155

Rutter, M., & Yule, W. (1975). The concept of specific reading retardation. *Journal of child psychology and psychiatry, and allied disciplines, 16*(3), 181–197. https://doi.org/10.1111/j.1469-7610.1975.tb01269.x

Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods, 8*(2), 597-599. https://doi.org/10.22237/JMASM/1257035100

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X. N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral cortex (New York, N.Y. : 1991), 28*(9), 3095–3114. https://doi.org/10.1093/cercor/bhx179

Specht K. (2020). Current Challenges in Translational and Clinical fMRI and Future Directions. *Frontiers in psychiatry, 10*, 924. https://doi.org/10.3389/fpsyt.2019.00924

Swagerman, S. C., van Bergen, E., Dolan, C., de Geus, E., Koenis, M., Hulshoff Pol, H. E., & Boomsma, D. I. (2017). Genetic transmission of reading ability. *Brain and language, 172*, 3–8. https://doi.org/10.1016/j.bandl.2015.07.008

Torgesen, J.K., Wagner, R.K., & Rashotte, C.A. (2012). Test of word reading efficiency–2nd ed. (TOWRE-2). Austin, TX: Pro-Ed.

Turner, B. O., Paul, E. J., Miller, M. B., & Barbey, A. K. (2018). Small sample sizes reduce the replicability of task-based fMRI studies. *Communications biology, 1*, 62. https://doi.org/10.1038/s42003-018-0073-z

van Bergen, E., de Jong, P. F., Maassen, B., Krikhaar, E., Plakas, A., & van der Leij, A. (2014). IQ of four-year-olds who go on to develop dyslexia. *Journal of learning disabilities, 47*(5), 475–484. https://doi.org/10.1177/0022219413479673

van Dijk, K. R., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., & Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *Journal of neurophysiology, 103*(1), 297–321. https://doi.org/10.1152/jn.00783.2009

Wang, Q., Hu, K., Wang, M., Zhao, Y., Liu, Y., Fan, L., & Liu, B. (2021). Predicting brain age during typical and atypical development based on structural and functional neuroimaging. *Human Brain Mapping, 42*(18), 5943–5955. https://doi.org/10.1002/hbm.25660

Wang, J., Ren, Y., Hu, X., Nguyen, V. T., Guo, L., Han, J., & Guo, C. C. (2017). Test-retest reliability of functional connectivity networks during naturalistic fMRI paradigms. *Human brain mapping, 38*(4), 2226–2241. https://doi.org/10.1002/hbm.23517

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain connectivity, 2*(3), 125–141. https://doi.org/10.1089/brain.2012.0073

Xia, Z., Hoeft, F., Zhang, L., & Shu, H. (2016). Neuroanatomical anomalies of dyslexia: Disambiguating the effects of disorder, performance, and maturation. *Neuropsychologia, 81*, 68–78. https://doi.org/10.1016/j.neuropsychologia.2015.12.003

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of neurophysiology, 106(3), 1125–1165. https://doi.org/10.1152/jn.00338.2011

Supplementary Material

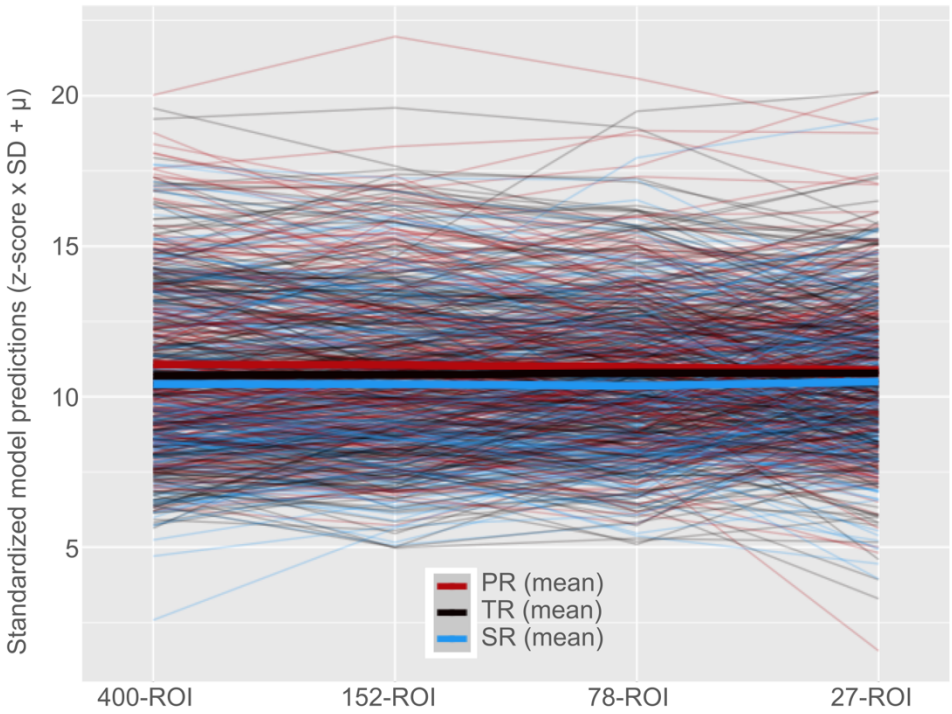
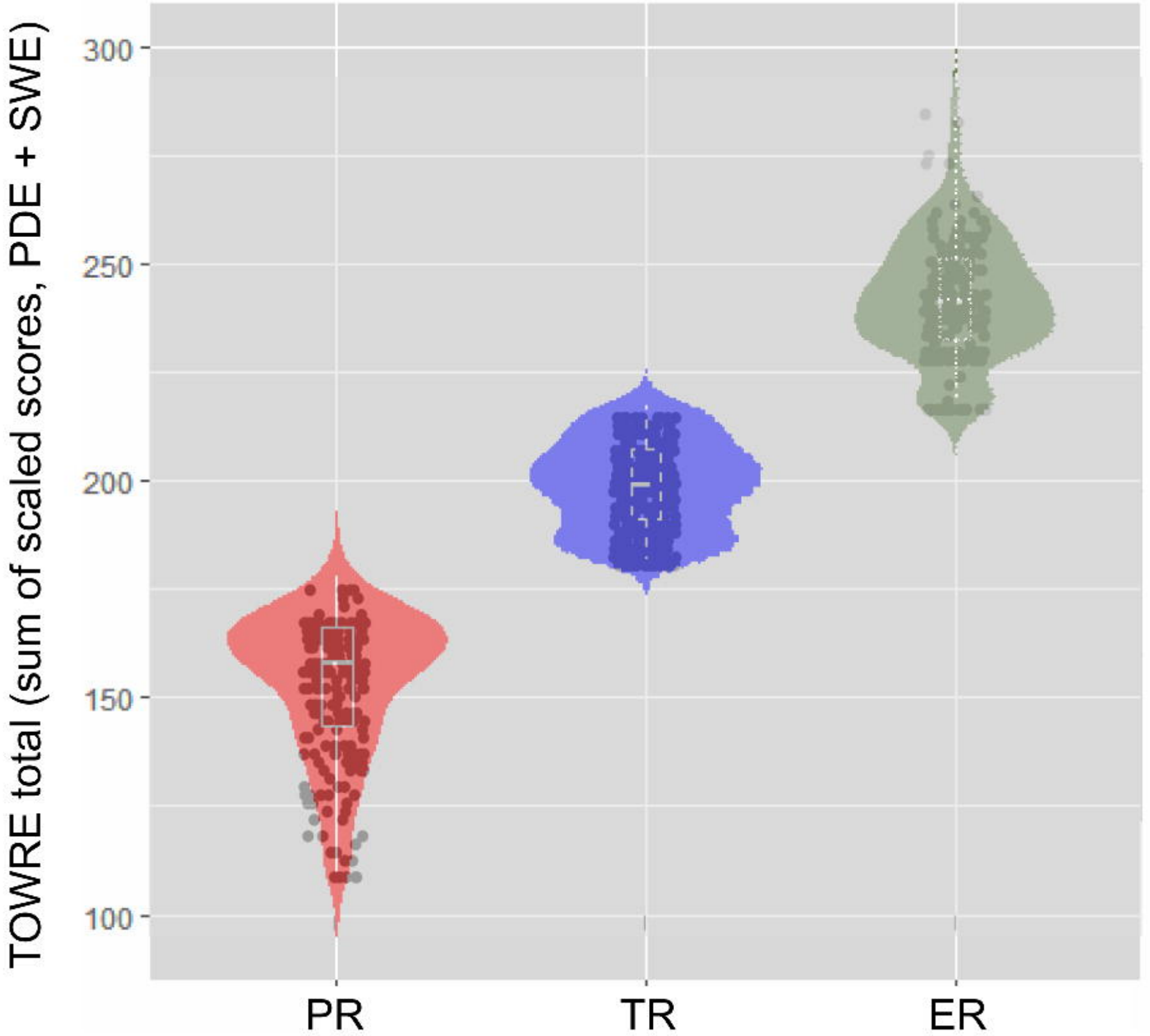
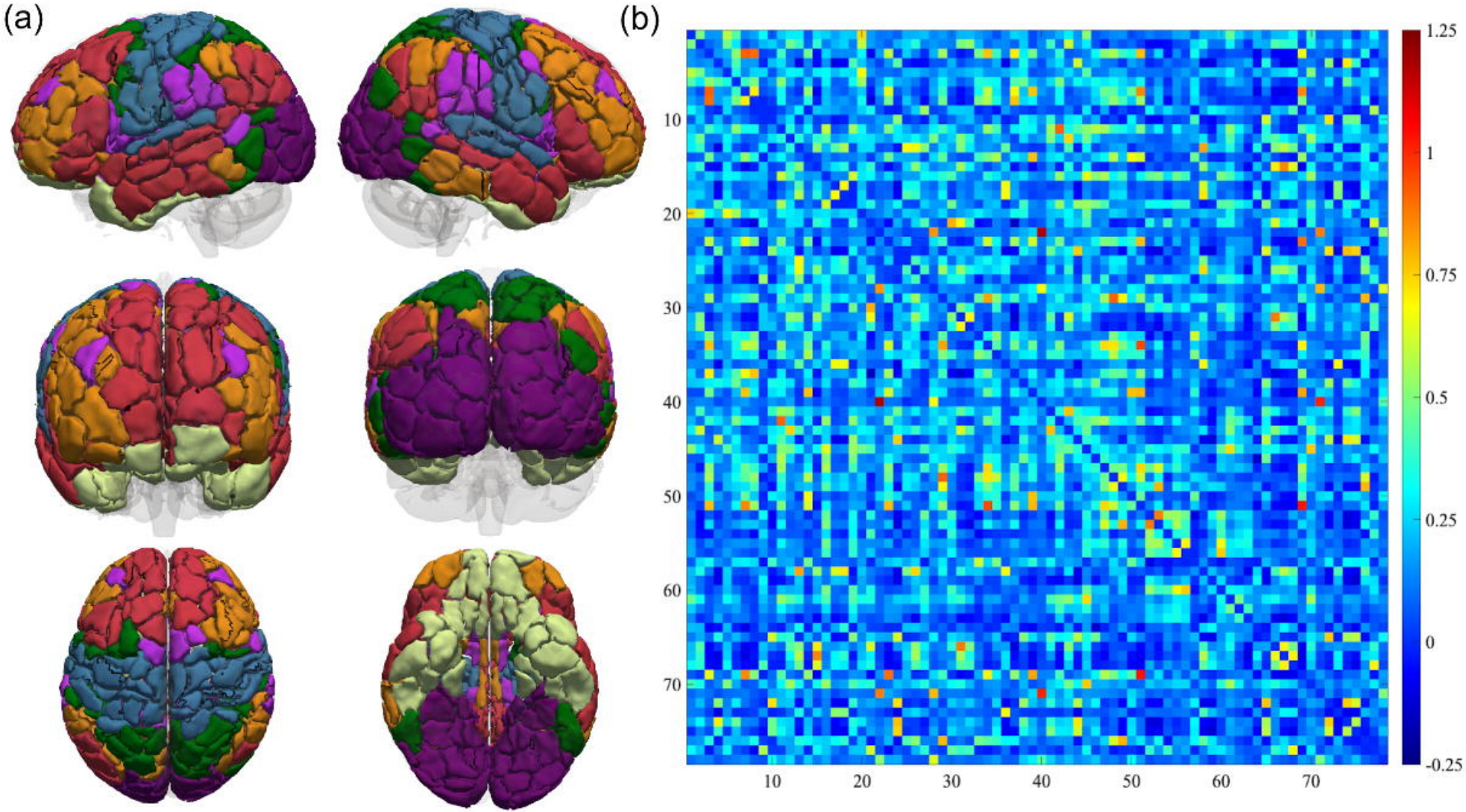
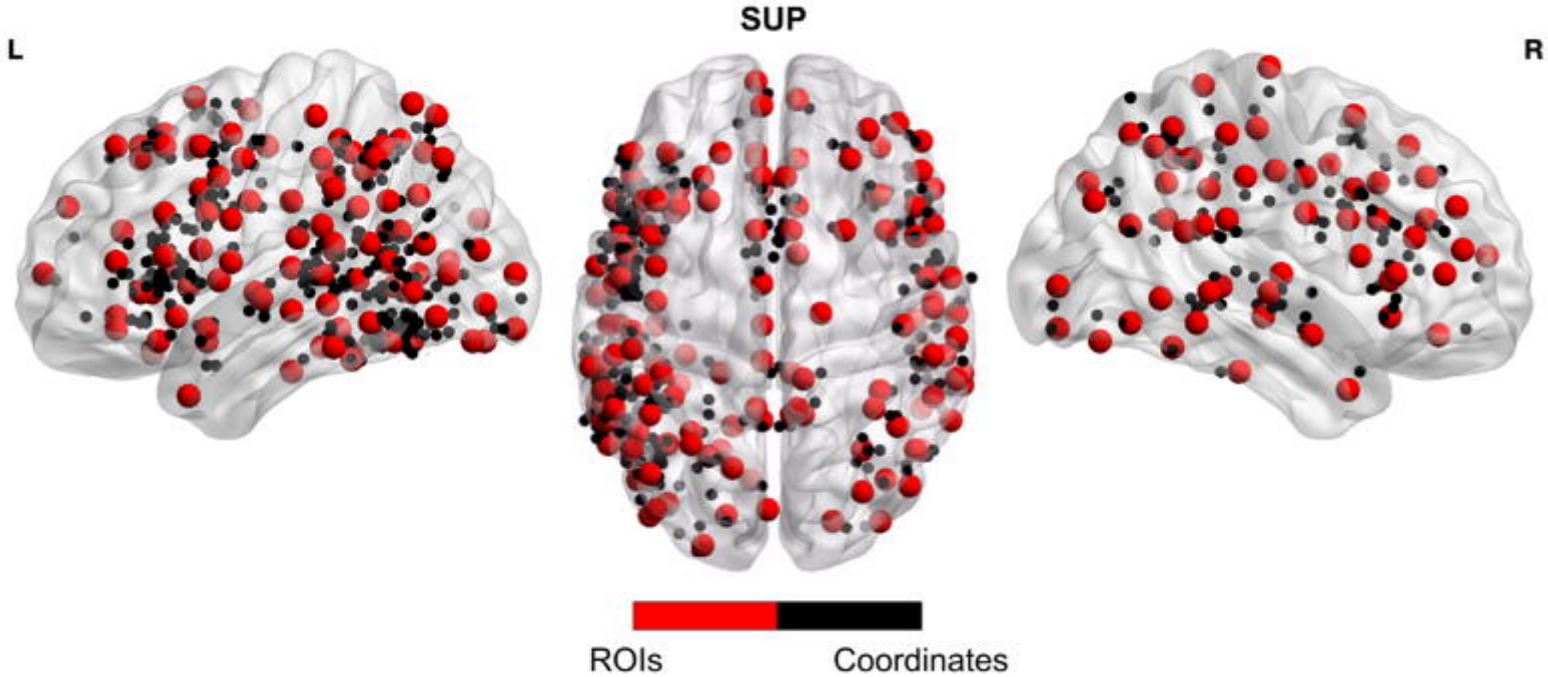
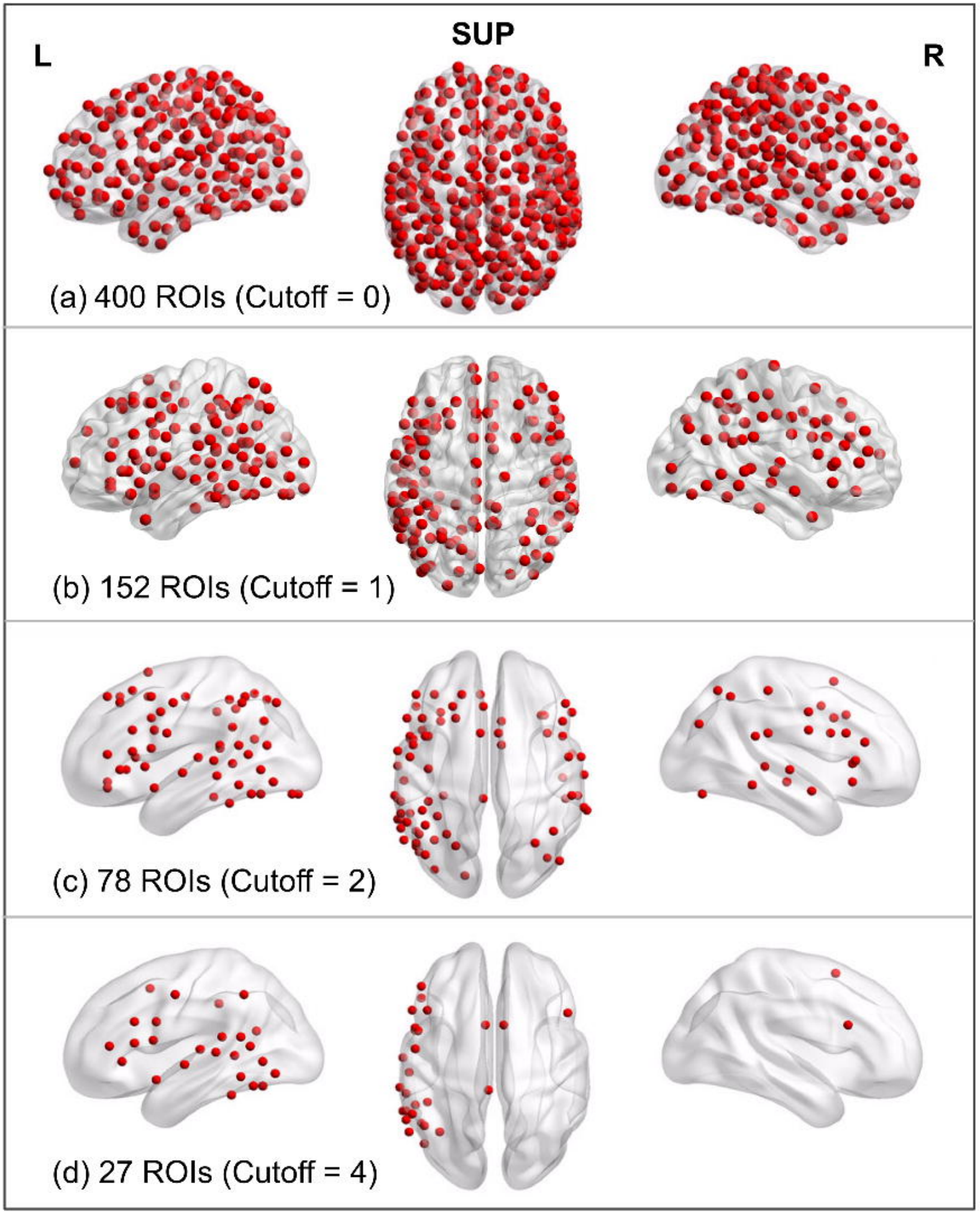


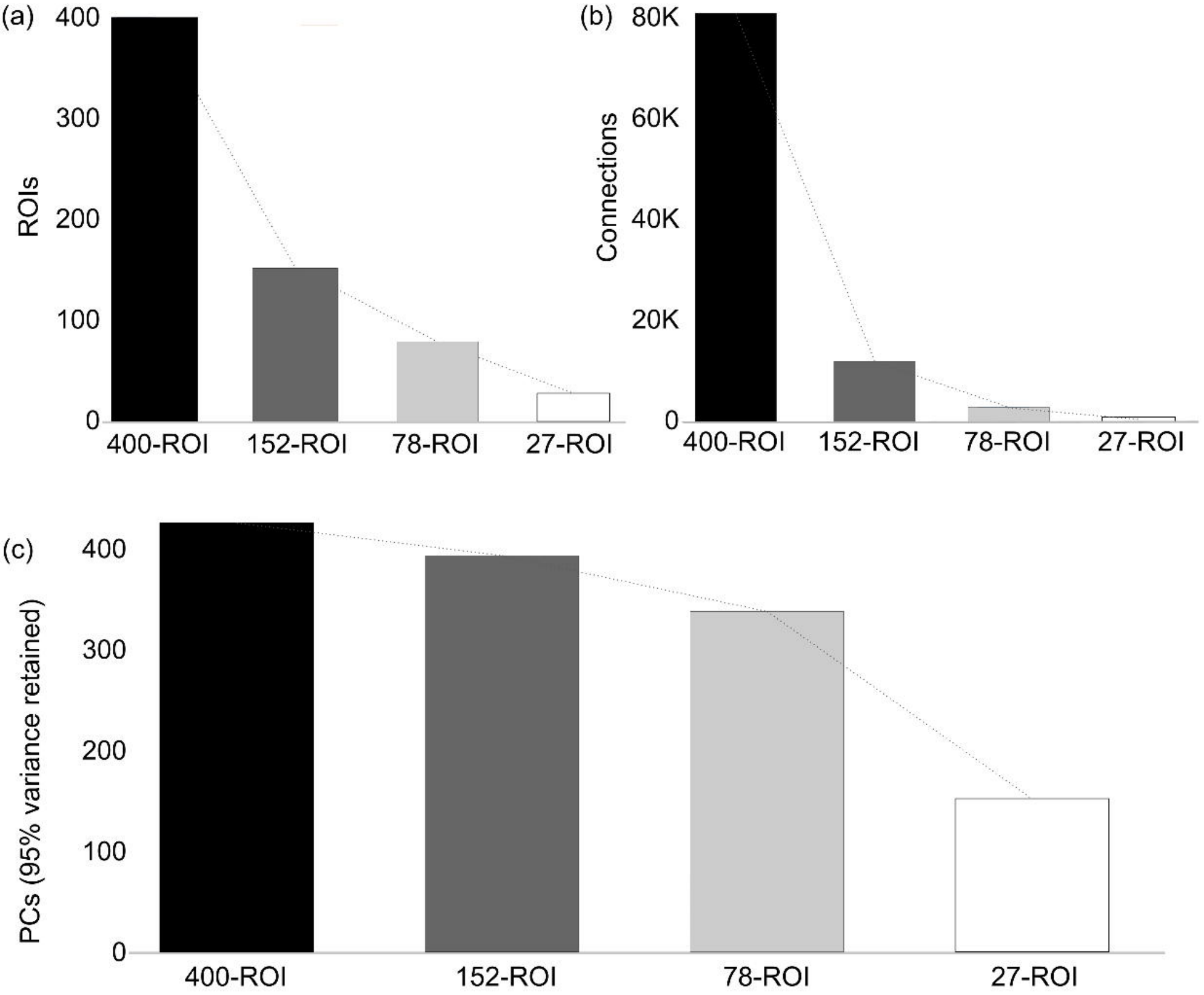
Figure S1. Individual line plots for each participant. Each connected line running from left-right shows the average prediction for a given participant across all ROI models. Thicker lines represent the group average predictions for each model.

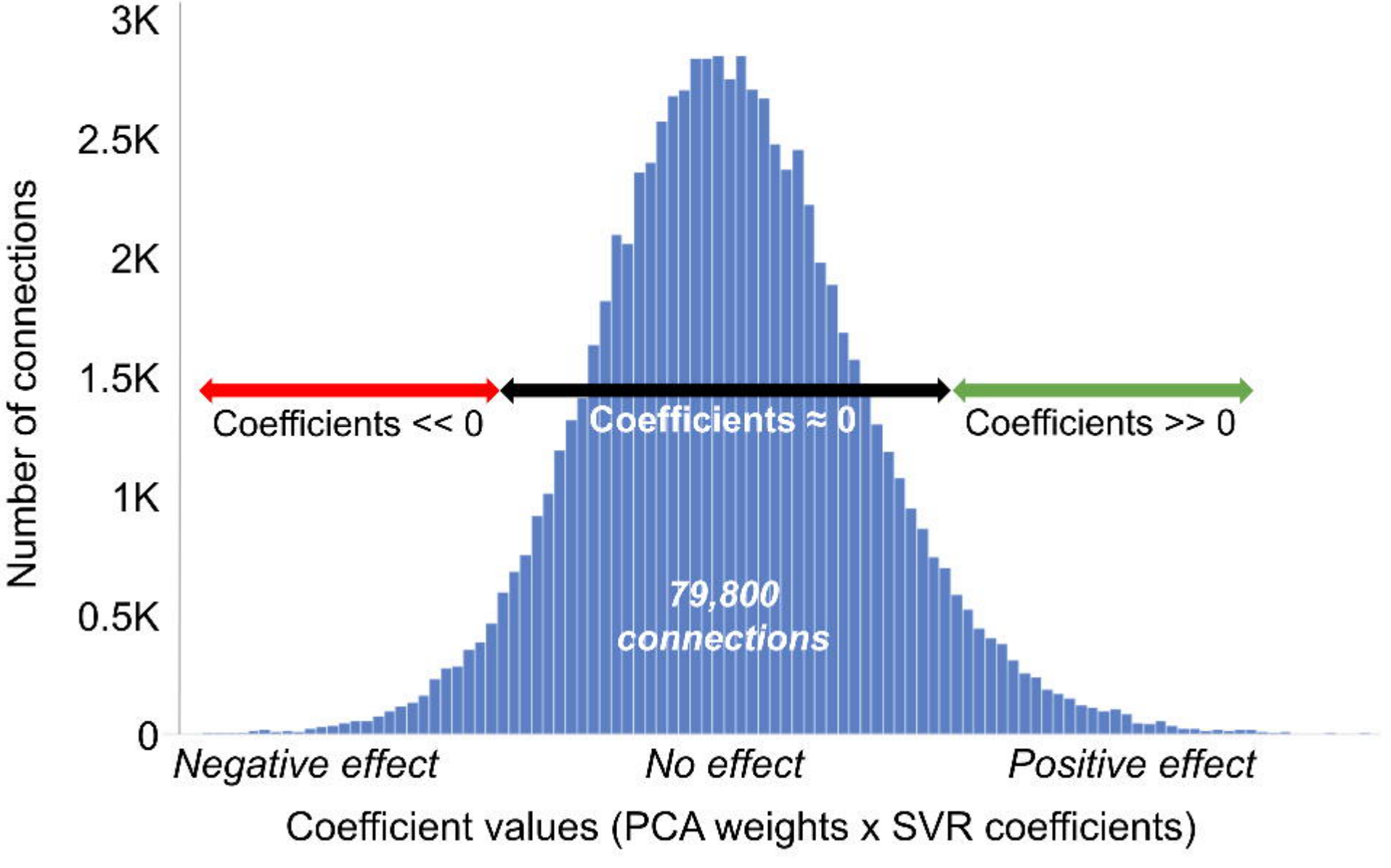

TOWRE total (sum of scaled scores, PDE + SWE)
300
250
200
150
100
PR
TR
ER

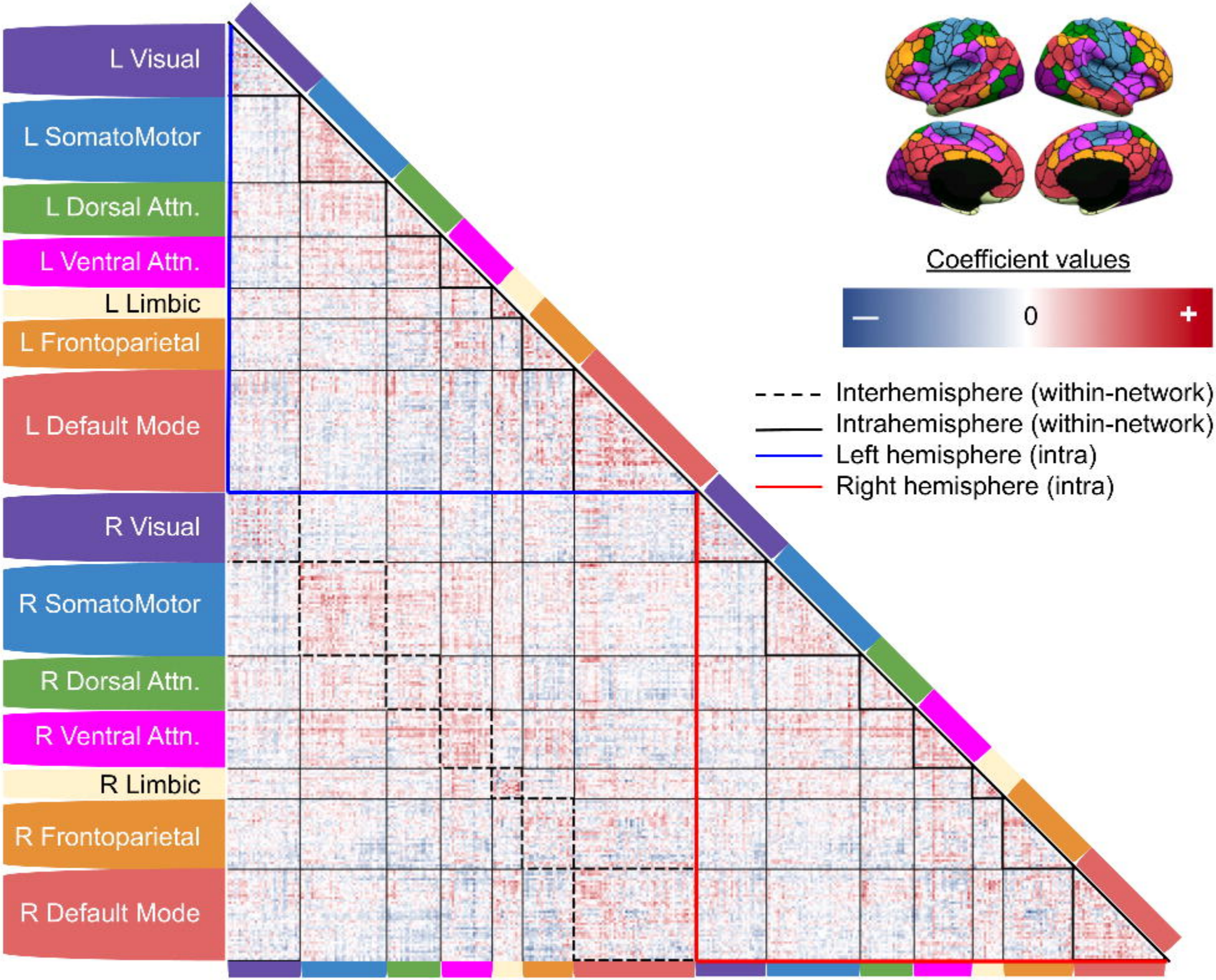

(a)
(b)
10
20
30
40
50
60
70
1.25
1
0.75
0.5
0.25
0
-0.25

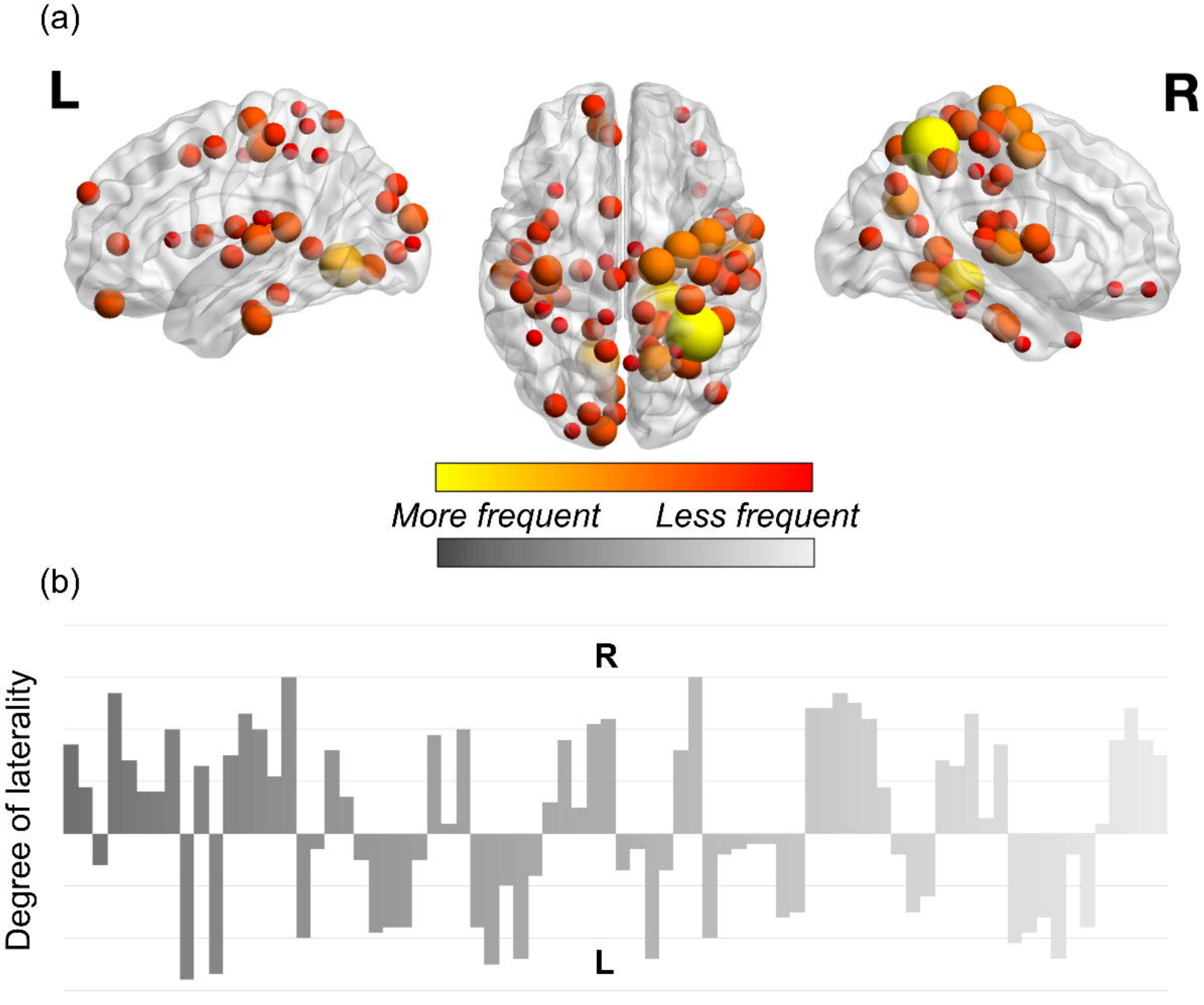

SUP
L
R
ROIs
Coordinates

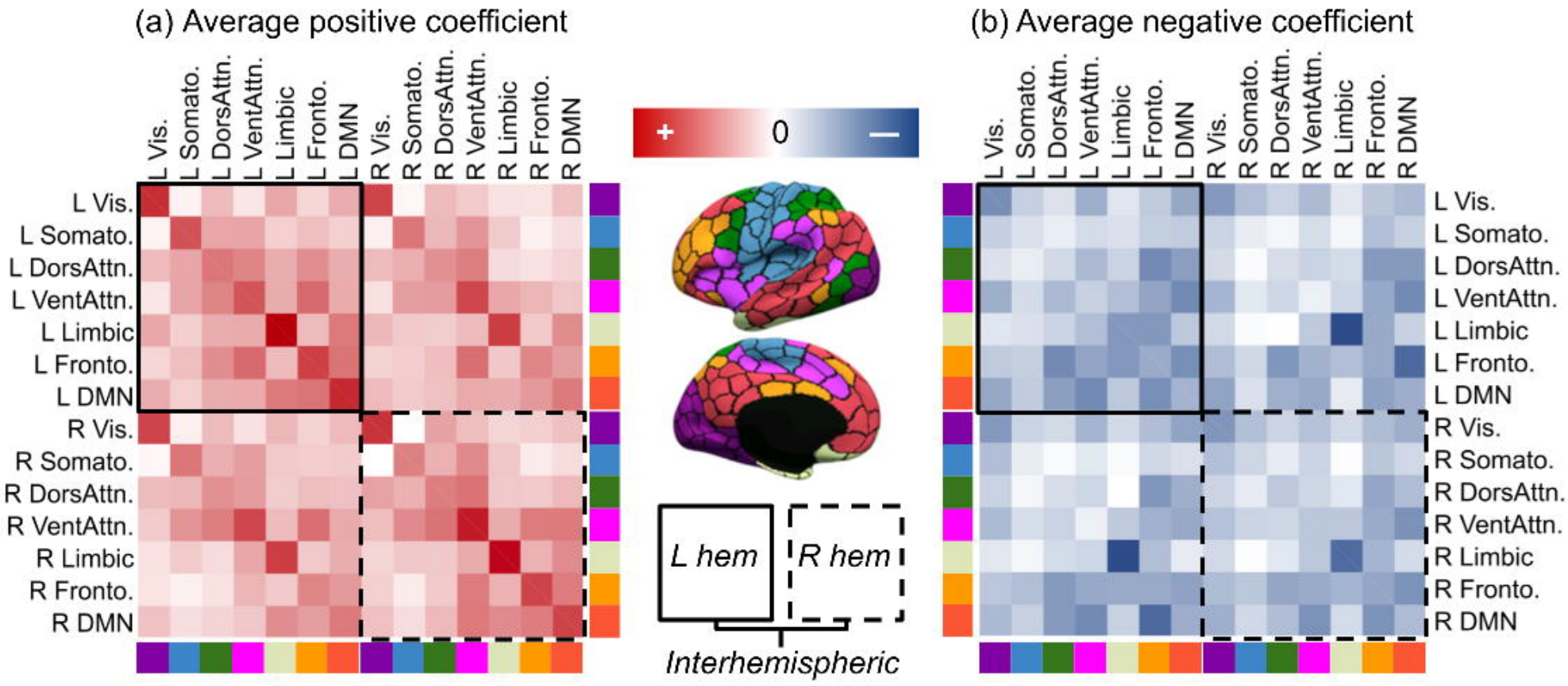

SUP
L
R
(a) 400 ROIs (Cutoff = 0)
(b) 152 ROIs (Cutoff = 1)
(c) 78 ROIs (Cutoff = 2)
(d) 27 ROIs (Cutoff = 4)

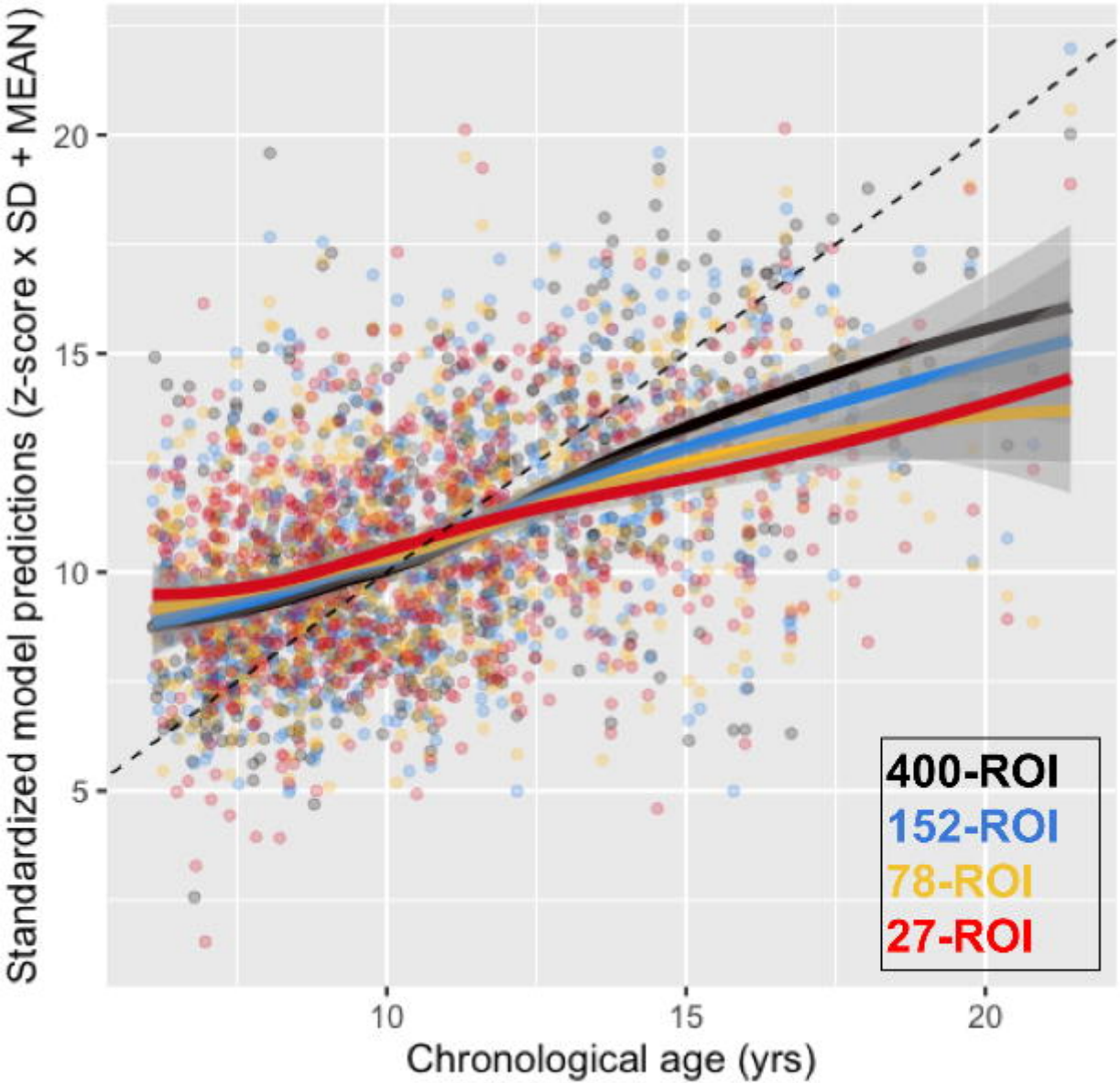

(a)
ROIs
400
300
200
100
0
400-ROI
152-ROI
78-ROI
27-ROI
(b)
Connections
80K
60K
40K
20K
0
400-ROI
152-ROI
78-ROI
27-ROI
(c)
PCs (95% variance retained)
400
300
200
100
0
400-ROI
152-ROI
78-ROI
27-ROI

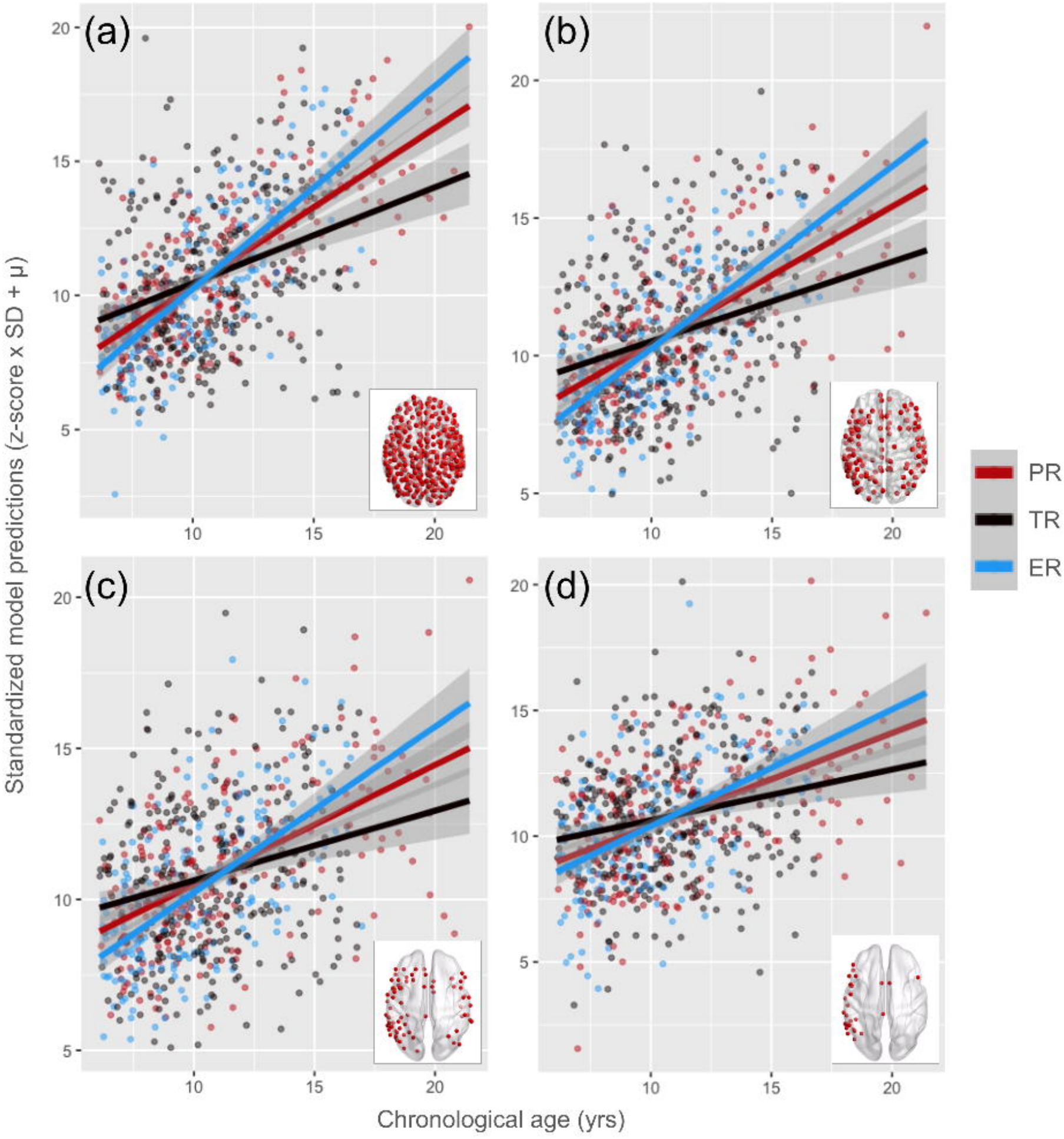

3K
2.5K
2K
1.5K
1K
0.5K
0
Number of connections
Coefficients << 0
Coefficients ≈ 0
Coefficients >> 0
79,800
connections
Negative effect
No effect
Positive effect
Coefficient values (PCA weights x SVR coefficients)

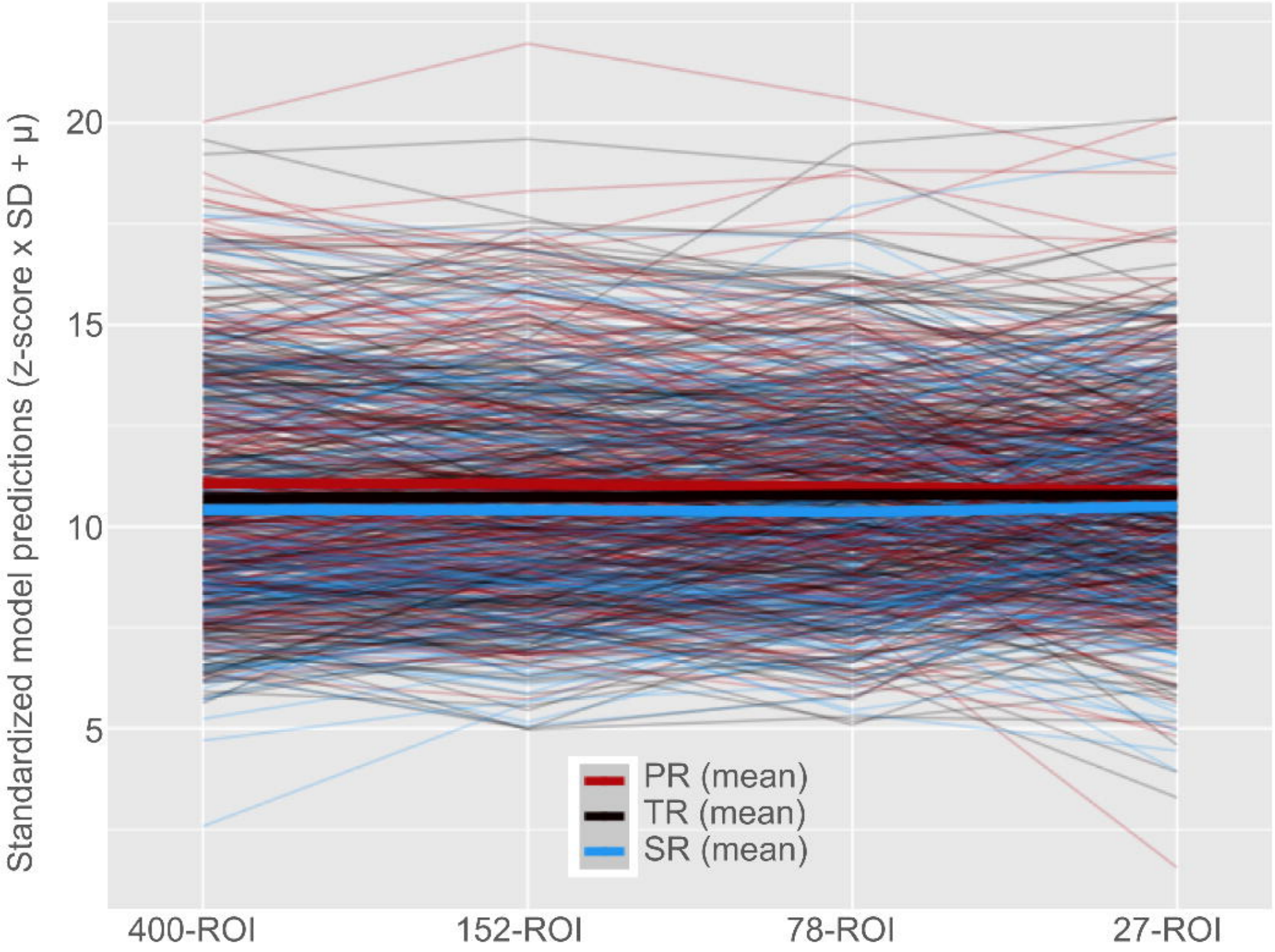

L Visual
L SomatoMotor
L Dorsal Attn.
L Ventral Attn.
L Limbic
L Frontoparietal
L Default Mode
R Visual
R SomatoMotor
R Dorsal Attn.
R Ventral Attn.
R Limbic
R Frontoparietal
R Default Mode
Coefficient values
–
0
+
Interhemisphere (within-network)
Intrahemisphere (within-network)
Left hemisphere (intra)
Right hemisphere (intra)


(a)
L
R
More frequent
Less frequent
(b)
R
Degree of laterality
L


(a) Average positive coefficient
(b) Average negative coefficient
L Vis.
L Somato.
L DorsAttn.
L VentAttn.
L Limbic
L Fronto.
L DMN
R Vis.
R Somato.
R DorsAttn.
R VentAttn.
R Limbic
R Fronto.
R DMN
+
0
−
L hem
R hem
Interhemispheric


Standardized model predictions (z-score x SD + MEAN)
20
15
10
5
10
15
20
Chronological age (yrs)
400-ROI
152-ROI
78-ROI
27-ROI


(a)
(b)
(c)
(d)
Standardized model predictions (z-score x SD + μ)
Chronological age (yrs)
PR
TR
ER


Standardized model predictions (z-score x SD + μ)
20
15
10
5
PR (mean)
TR (mean)
SR (mean)
400-ROI
152-ROI
78-ROI
27-ROI