Impact of brain parcellation on prediction error in models of cognition and demographics

ABSTRACT

Brain connectivity analysis begins with the selection of a parcellation scheme that will define brain regions as nodes whose connections will be studied. Brain connectivity has already been used in predictive modelling of cognition, but it remains unclear if parcellation schemes can systematically impact the predictive model error. In this work, structural, functional and combined connectivity were each defined with 5 parcellation schemes. Principal Component Regression with elastic net regularisation was fitted to each connectivity to predict individual differences in age, education, sex, Executive Function, Self-regulation, Language, Encoding and Sequence Processing. It was found that low-resolution functional parcellation consistently performed above chance at producing generalisable models of both demographics and cognition. However, no single parcellation scheme proved superior at predictive modelling. In addition, although parcellation schemes have influenced the global organisation of each connectivity type, this difference could not account for the error of the models. Taken together, this demonstrates that while high-resolution parcellations may be beneficial for modelling specific individual differences, partial voluming of signals achieved by higher resolution of parcellation likely disrupts model generalisability.

INTRODUCTION

Neuroimaging research has demonstrated that adaptive behaviour relies not only on localised activation in the brain but also on effective coordination of activity across remote neuronal populations (Alvarez & Squire, 1994; Seeley et al., 2007). To understand how information is exchanged in the human brain some authors have studied the white matter connections between regions (structural connectivity, SC) and statistical associations between their activity (functional connectivity, FC) (Park & Friston, 2013). The integrity of both SC and FC have been related to healthy cognitive function throughout the lifespan (Salami et al., 2014), whereas disruption to their integrity has been demonstrated to characterise many psychiatric, developmental and clinical diagnostics (de Kwaasteniet et al., 2013; Guye et al., 2010; Hahn et al., 2013; Keller et al., 2007; van den Heuvel & Fornito, 2014). Consequently, the study of the relationship between brain connectivity and cognition has become increasingly popular (Farahani et al., 2019; Hoppenbrouwers et al., 2014; Manca et al., 2018).

One of the most central questions in the analysis of brain connectivity is how the human brain is parcellated into brain regions (aka nodes) (Eickhoff et al., 2018; Yao et al., 2015). In the analysis of brain connectivity, brain regions constitute whole units whose patterns of connections are studied, and these patterns of connections are often used to explain and predict cognitive function. Therefore, brain parcellation is the foundation of

connectivity analysis. Nearly two dozen different brain parcellations are now available (Arslan et al., 2018; Dickie et al., 2017; Eickhoff et al., 2018; Lawrence et al., 2021). Each parcellation has been defined using different neuroimaging data and parcel generation algorithms. Consequently, each parcellation offers different node locations, sizes and shapes. However, clinical and cognitive neuroimaging research does not yet have a standard or most common parcellation. In addition, it is unlikely that a 'one size fits all' standard will emerge because every parcellation method has been defined to reflect specific properties of the brain (Eickhoff et al., 2018). To illustrate, we may choose to parcellate the brain based on local information histological boundaries (Amunts & Zilles, 2015; Brodmann, 1909) or clustering of a group of similarly active voxels (Cohen et al., 2008).

Despite the lack of an accepted standard, parcellation choice is an important decision in network analysis and it carries profound consequences for research. It has been empirically demonstrated that parcellation can impact the organisation of brain networks. For example, evidence demonstrates that the parcellation method impacts the small world coefficient of SC and FC, which refers to the network's tendency to segregate into clusters of strongly connected nodes relative to its tendency to produce short paths between any pair of nodes (Fornito, 2010; Wang et al., 2009; Zalesky et al., 2010). In addition, network parcellation also impacts their probability to produce highly connected nodes (aka degree distribution) (Fornito, 2010; Wang et al., 2009; Zalesky et al., 2010). Furthermore, the choice of parcellation resolution has also been demonstrated to impact the similarity of structural and functional connectivity (Akiki & Abdallah, 2019; Ashourvan et al., 2019; Honey et al., 2009). Finally, parcellation choice alone has been demonstrated to impact the explanatory and predictive power of neuroimaging models of cognition in healthy and clinical populations (Dhamala et al., 2021; Mellema et al., 2020; Ota et al., 2014; Pervaiz et al., 2020). In particular, some evidence suggests that high-resolution parcellation may benefit the predictive power of models of cognition (Mellema et al., 2020; Pervaiz et al., 2020).

However, no work has yet been conducted to assess if the parcellation method systematically influences network architecture and the explanatory and/or predictive power of models of cognition. The present work investigated this question by defining SC and FC with a variety of atlases - the resting state FC-defined atlas with either 100, 200 and 300 parcels (Shen et al., 2013), anatomically-defined atlas [(E. T. Rolls et al., 2020)] and a hybrid of structural and functional connectivity atlas (Xia et al., 2013). Our work then produced connectivity-based models of cognition and used graph theory to assess the global differences in network anatomy that may explain model differences. Doing so, we tested the hypotheses that (1) SC models of cognition will be more effective when connectivity is defined with an anatomy-based parcellation (AAL3) than parcellation based on a functional activation (Shen et al., 2013) and vice versa will apply to FC, (2) use of high-resolution parcellation will improve explanatory and predictive power models of cognition regardless of connectivity type, (3) differences in model explanatory and predictive power will systematically map with differences in network organisation, as measured by graph theory.

METHODS

Participants

Neuroimaging and cognitive data were obtained for 250 unrelated subjects from the 1200-subject release of the Human Connectome Project (HCP). For consistent treatment of behavioural and neuroimaging subjects' data selection, one subject was excluded from the neuroimaging analysis due to incomplete behavioural data. The sample consisted of 138 females and 111 males in the age range of 22 and 36 years.

Measures of cognition

The present work used the measures of cognition from our previous work (Litwińczuk et al., 2022). Briefly, Principal Component Analysis (PCA) with VARIMAX rotation was used as a feature extraction method from the behavioural dataset. Analysed tasks included: Picture Sequence Memory, Dimensional Change Card Sort, Flanker Inhibitory Control and Attention Task, Penn Progressive Matrices, Oral Reading Recognition, Picture Vocabulary, Pattern Comparison Processing Speed, Delay Discounting, Variable Short Penn Line Orientation Test, Short Penn Continuous Performance Test, Penn Word Memory Test, and List Sorting. These assessments were obtained from the Blueprint for Neuroscience Research–funded NIH Toolbox for Assessment of Neurological and Behavioral function (http://www.nihtoolbox.org) and tasks from the Penn computerized neurocognitive battery (Gur et al., 2010). The extracted PCA rotated components reflected specific latent cognitive domains, interpreted as Executive Function, Self-regulation, Language, Encoding and Sequence Processing. The present work uses the PCA scores obtained previously for each cognitive domain.

Minimally processed neuroimaging data

The HCP provides minimally processed neuroimaging data that was used here, the data acquisition and processing pipeline has been discussed in detail by (Glasser et al., 2013). All neuroimaging data were collected with a 3T Siemens "Connectome Skyra" scanner that uses the Siemens 32-channel RF receive head coil and with an SC72 gradient insert (Ugurbil et al., 2013). Here, we utilised Version 3 of the minimal processing pipeline implemented with FSL 5.0.6 (Jenkinson et al., 2012) and FreeSurfer 5.3.0-HCP (Dale et al., 1999).

T1 weighted MR images were acquired with a 3D MPRAGE sequence (TR = 2400 ms, TE = 2.14, TI = 1000 ms, flip angle = 8°, FOV = 224 by 224 mm, voxel size = 0.7 mm isotropic). Rs-fMRI data was collected using the gradient-echo EPI (TR = 720 ms, TE = 33.1 ms, flip angle = 52°, FOV = 208 by 180 mm, 70 slices, thickness = 2.0 mm, size = 2.0 mm isotropic). Scans were collected in two sessions, each lasting approximately 15 minutes. The rs-fMRI data were collected both in left-to-right and right-to-left directions. In addition, in the original data, spin echo phase reversed images were acquired for registration with T1 images and the spin echo field maps were acquired for bias field correction. Diffusion-weighted MR images were acquired with spin-echo EPI sequence (TR = 5520 ms, TE = 89.5 ms, flip angle = 78°, refocusing flip angle = 160°, FOV = 210 by 180 mm, 111 slices, thickness = 1.25 mm, size = 1.25 mm

isotropic). Each gradient consisted of 90 diffusion weighting directions plus 6 b=0. There were 3 diffusion-weighed shells of b=1000, 2000, and 3000 s/mm^2. SENSE1 multi-channel image reconstruction was used (Sotiropoulos et al., 2013).

Additional processing of neuroimaging data

Neuroimaging data were processed following the same processing pipeline as in our previous work (Litwińczuk et al., 2022).

Structural data and structural connectivity calculation

As additional steps to the minimal processing pipeline, the diffusion data were further analysed using the BEDPOSTX procedure in FSL, which runs Markov Chain Monte Carlo sampling to estimate probability distributions on diffusion parameters at each voxel. This information was used in the FDT module of FSL to run ROI-to-ROI probabilistic tractography with ProbtrackX. Tractography was run between parcels obtained with a high-resolution functionally defined brain parcellation with 278 parcels (Shen et al., 2013). During tractography, 5000 streamlines were initiated from each voxel with a step length of 0.5 mm (Behrens et al., 2007; Behrens et al., 2003; Jenkinson et al., 2012). Streamlines were constrained with a curvature threshold of 0.2, a maximum of 2000 steps per streamline and a volume fraction threshold of subsidiary fiber orientations of 0.01. An SC matrix between regions was constructed by first counting the number of streamlines originating from a seed region i that reached a target region j (M_{ij}). These counts are asymmetric since the count of streamlines from region i to j is not necessarily equal to the count of streamlines from region j to i ($M_{ij} \neq M_{ji}$), but they are highly correlated for all subjects (lowest Pearson's Correlation was 0.76, $p < 0.001$). Based on these counts, the weight W_{ij} (entries of the SC matrix) between any two pairs of regions i and j was defined as the ratio of the total streamline counts in both directions ($M_{ij} + M_{ji}$), to the maximum possible number of streamlines that can be shared between the two regions, which is $(N_i + N_j) * 5000$ (where N_i and N_j are the numbers of seed voxels in regions i and j, respectively):,

$$W_{ij} = \frac{(M_{ij} + M_{ji})}{(N_i + N_j) * 5000}$$

Similar to previous studies, the weight W_{ij} can be interpreted as capturing the connection density (number of streamlines per unit surface) between nodes i and j, which accounts for possible bias due to the different sizes of the seed regions (Friston et al., 2008; Ingalhalikar et al., 2013). Note that the SC matrix defined based on these weights is symmetric because swapping around the regions' indices does not change the result; and it is also normalised between 0 and 1, because the maximum value of the numerator can only be reached when all streamlines originating from each of region reach the other region, so that $M_{ij} = N_i * 5000$ and $M_{ji} = N_j * 5000$, which gives $W_{ij} = 1$. Evidence suggests that structural connectivity is most sensitive to individual differences with moderate-to-high

thresholding (Buchanan et al., 2020) and produces the least false positive and negative results (de Reus & van den Heuvel, 2013), therefore an 80% proportional threshold was applied.

Functional data and functional connectivity calculation

The minimally processed images were obtained for rs-fMRI to compute FC based on pair-wise correlations (Glasser et al., 2013). Next, the following steps were taken to further process data using the CONN Toolbox (Whitfield-Gabrieli & Nieto-Castanon, 2012) with the use of the standard FC processing pipeline (Nieto-Castanon, 2020). Briefly, images were realigned, a slice-timing correction was conducted, and outlier detection of functional images for scrubbing was performed with Artefact Detection Tools (ART, https://www.nitrc.org/projects/artifact_detect/). Grey matter, white matter, cerebrospinal fluid and non-brain tissues were then segmented. Images were normalized and smoothed with a 6 mm Full Width at Half Maximum Gaussian kernel. Next, the data was denoised with default Conn denoising options using the anatomical component-based noise correction procedure (Behzadi et al., 2007). This procedure removes artefactual components from the data, including noise components from cerebral white matter and cerebrospinal areas, subject-motion parameters (Friston et al., 1996), identified outlier scans (Power et al., 2014), and constant and first-order linear session effects (Whitfield-Gabrieli & Nieto-Castanon, 2012). Then standard denoising steps were applied including scrubbing, motion regression and application of a high pass filter (0.01 Hz cut-off), and a low pass filter (0.10 Hz cut-off).

FC analysis was performed based on the same high-resolution brain parcellation used in the SC computations (Shen et al., 2013). The average blood oxygenation level-dependent signal in each ROI was obtained and the pairwise (ROI-to-ROI) correlation of the averaged signals was calculated. Since the CONN toolbox produces Fisher's Z-scores (Fisher, 1915), a hyperbolic tangent function was used to reverse Fisher's transformation and obtain original correlation values ranging between -1 and 1. Negative correlations were transformed to positive by taking their absolute values and a proportional 80% FC threshold was then applied (Garrison et al., 2015; van den Heuvel et al., 2017).

Model construction

Models of 5 cognitive domains were separately constructed using FC, SC or a combination of SC and FC referred to as combined connectivity (CC). Network nodes were defined with 5 brain parcellations:

1. Low-resolution functional parcellation (Shen 100) (Shen et al., 2013)
2. Moderate-resolution functional parcellation (Shen 200) (Shen et al., 2013)
3. High-resolution functional parcellation (Shen 300) (Shen et al., 2013)
4. Automated Anatomical Labelling (AAL3) [(Edmund T. Rolls et al., 2020)]
5. Brainnetome (Xia et al., 2013).

Independent models were constructed for each connectivity type and with each parcellation scheme. The Principal Component Regression approach was followed with the elastic net regularisation of beta coefficients:

1) The confounding effect of age, gender and education was regressed out of response variable (*y*)
2) Principal Components Analysis (PCA) was conducted to obtain orthogonal components of the connectivity matrix (X)
3) A regression model was fitted with elastic net regularisation
4) Beta values obtained in PCA space were projected back to the original connectome space and used to produce fitted responses.

To tune the values of alpha and lambda, 5-fold cross-validation was repeated 50 times (repeated CVT). During repeated CVT, the whole sample was divided into a training sample (80% of the total sample) and a validation sample (20% of the total sample). The PCR model was fitted to the training sample. The validation sample was used to assess the out-of-sample model performance. Optimal alpha and lambda values were determined as the values that produced the model with the highest median coefficient of determination between the validation sample's observed responses and predicted responses.

In addition, cross-validation has been demonstrated to produce a combination of pessimistic and optimistic biases of model quality. A pessimistic bias can be observed due to the reduced size of the training sample used during the cross-validation procedure. An optimistic bias can be observed due to selecting the model configuration as the best among many tries. To correct this, Bootstrap Bias Corrected Cross-Validation has been implemented (Tsamardinos et al., 2018). During this procedure, out-of-sample predictions generated during repeated CVT and their corresponding observed outcomes were bootstrapped. Bootstrapped samples were used to conduct a configuration selection strategy, which returns the best-performing alpha and lambda terms according to minimum average loss. The not-bootstrapped samples are used to assess the loss of selected terms. This was repeated 5,000 times. The explained variation (coefficient of determination) was used as an estimate of the loss (Poldrack et al., 2020). The resultant population of bootstrapped statistics was used to produce mean performance estimates of the PCR learning strategy and corresponding confidence intervals.

Finally, the permutation (randomisation) test was used to assess how likely it is to get the observed models' performance by chance. Specifically, the out-of-sample predictions of the optimal alpha and lambda terms produced by repeated CVT were randomised (sampled without replacement) 5,000 times and the models' performance statistics (coefficient of determination) were estimated for each randomisation. This null distribution was then used to assess the observed model performance statistics in the non-permuted data. That is, the

probability for testing models' performance was determined by computing the proportion of resampled statistics at least as high or greater than the observed statistics.

Model comparison

All following analysis was done separately for each cognitive domain and connectivity type, and the analysis compared the effects of parcellation choice.

To compare the model whole-sample predictive power, Akaike Information Criterion (AIC) was obtained. The AIC balances the goodness-of-fit of the model (model accuracy) against its complexity (the number of parameters included in the model) so that reduced AIC is associated with improved model quality. Given any two models M_1 and M_2, a positive difference (ΔAIC = AIC(M_1)- AIC (M_2)) is interpreted as weak (barely worth a mention) (1-3 units), positive (3-20 units) or strong (20-150 units) evidence in favour of M_2 (Kass & Raftery, 1995). We used the AIC difference to compare the explanatory power of models defined with connectivity in each type of parcellation (functional or anatomical). Consequently, if two parcellations had similar explanatory power but one was coarser than the other, the coarser parcellation was favoured due to requiring fewer variables to model the same behaviour.

To compare model out-of-sample predictive power, the model out-of-sample predictions (generated during repeated CVT for the optimal alpha and lambda term) were bootstrapped 5000 times and the models' performance statistics were obtained. We used the non-parametric Wilcoxon rank sum tests for equal medians to assess the significance of differences in performance between different models. These comparisons were only done for models which performed better than chance, and the results were based on the coefficient of determination.

Graph theory

Graph theoretic measures were calculated based on the weighted, undirected SC and FC matrices of every subject, using The Brain Connectivity Toolbox (http://www.brain-connectivity-toolbox.net). Measures of global network organisation included small-world propensity, global efficiency, assortativity, modularity statistic, transitivity, and coreness statistic resulting from core-periphery partition. The clustering coefficient was obtained for small world propensity with Onnela's algorithm (Muldoon et al., 2016; Onnela et al., 2005). The shortest path length was calculated using the Floyd–Warshall Algorithm applied to the weighted graph obtained from the inverse of each connectivity matrix. Small-world propensity has been developed by Muldoon et al. (2016). Network modules for within module node degree and participation coefficient were defined with Newman's algorithm (Newman, 2006).

To explore how parcellation impacted global network organisation, a series of paired t-tests compared global graph theory measures across parcellation schemes in order to assess

if parcellation schemes impact network organisation. Then, a correlation analysis was conducted to establish if the overall network organisation, yielded by each parcellation scheme, relates to model error. Here, Pearson's correlation assessed associations between model residuals and global graph theory.

RESULTS

Demographics

AIC-based model comparison

Across all models of participant demographics, the benefit of defining SC, FC and CC using specific parcellations was barely worth a mention (ΔAIC< 3).

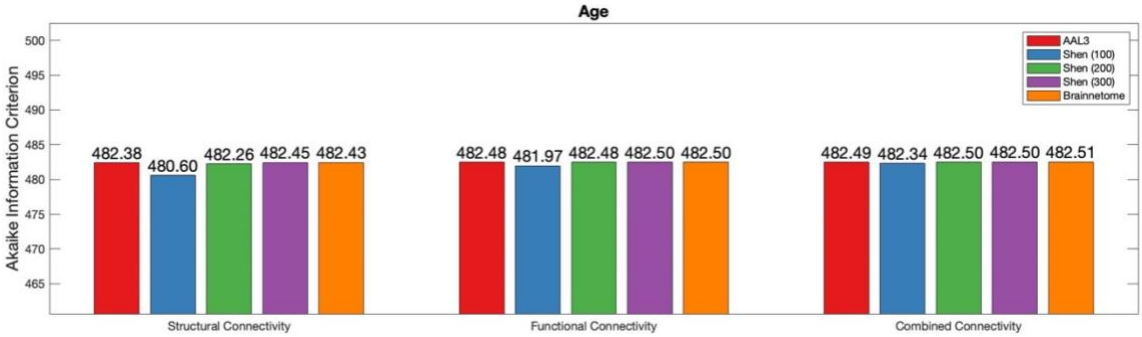

Figure 1 AIC model evidence for SC, FC and CC models of age, constructed with each parcellation scheme. Models with lower AIC values are favoured.

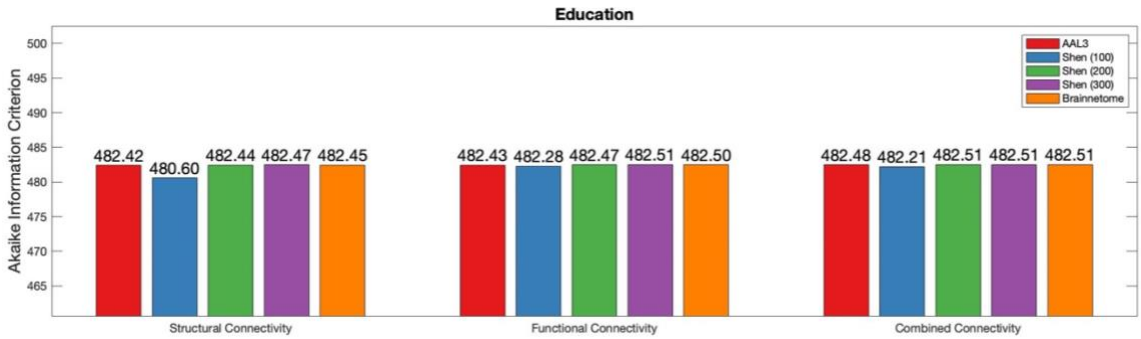

Figure 2 AIC model evidence for SC, FC and CC models of education, constructed with each parcellation scheme. Models with lower AIC values are favoured.

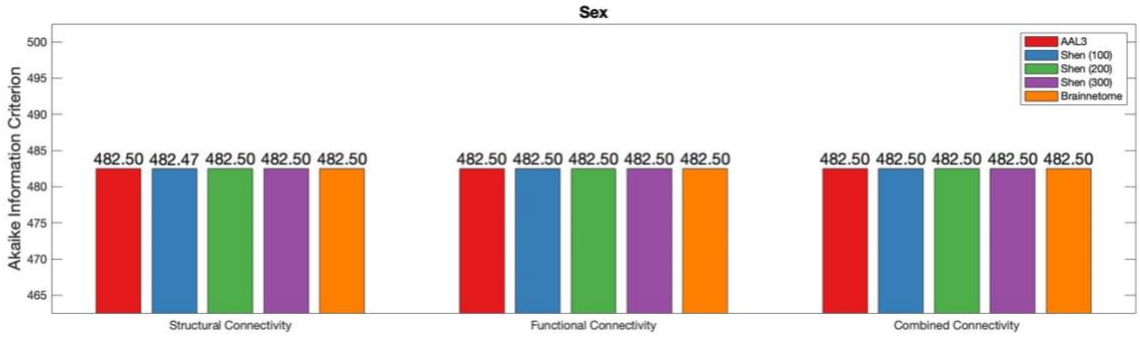

Figure 3 AIC model evidence for SC, FC and CC models of sex, constructed with each parcellation scheme. Models with lower AIC values are favoured.

Cross-validation-based model comparison

Results of the cross-validation of models of demographics are presented in Figures 4-6 and pairwise comparisons of the cross-validation errors are presented in Table 1. All following comparisons have passed the False Discovery Rate adjusted significance threshold of 0.0095 (Benjamini & Hochberg, 1995).

Cross-validation results have demonstrated that only functional connectivity defined using Shen 300 parcellation failed to produce out-of-sample predictions of Education. All remaining connectivity-based models performed above chance regardless of parcellation schemes used.

Age was most effectively predicted in the unseen sample using AAL3 parcellation in structural connectivity (vs Shen 100 Z = 14.07, vs Shen 200 Z = 3.76, vs Shen 300 Z = 18.47, vs Brainnetome Z = 9.27), Shen 300 in functional connectivity (vs AAL3 Z = 65.95, vs Shen 100 Z = 70.37, vs Shen 200 Z = 28.36, vs Brainnetome Z = 17.78) and Shen 200 in combined connectivity (vs AAL3 Z = 48.91, vs Shen 100 Z = 41.68, vs Shen 300 Z = 0.62, vs Brainnetome Z = 11.47).

Education was most effectively predicted in the unseen sample using AAL3 parcellation in structural connectivity (vs Shen 100 Z = 45.49, vs Shen 200 Z = 30.57, vs Shen 300 Z = 13.74, vs Brainnetome Z = 24.61). Shen 200 and Brainnetome parcellation performed equally well in functional connectivity (Z = 0.02). Shen 200 performed more effectively than AAL3 parcellation (Z = 52.78) and Shen 100 (Z = 11.91). Brainnetome parcellation performed more effectively than AAL3 parcellation (Z = 51.19) and Shen 100 (Z = 11.59). Brainnetome parcellation performed more effectively at predicting education than alternative parcellations in combined connectivity (vs AAL3 Z = 32.87, vs Shen 100 Z = 46.51, vs Shen 200 Z = 17.27, vs Shen 300 Z = 18.70).

Sex was most effectively predicted in an unseen sample using Shen 100 in structural connectivity (vs AAL3 Z = 15.09, vs Shen 200 Z = 3.45, vs Shen 300 Z = 5.98, vs Brainnetome Z

= 16.85), Shen 200 in functional connectivity (vs AAL3 Z = 63.69, vs Shen 100 Z = 30.34, vs Shen 200 Z = 66.43, vs Brainnetome Z = 6.52) and in combined connectivity (vs AAL3 Z = 33.98, vs Shen 100 Z = 14.08, vs Shen 300 Z = 18.19, vs Brainnetome Z = 9.53).

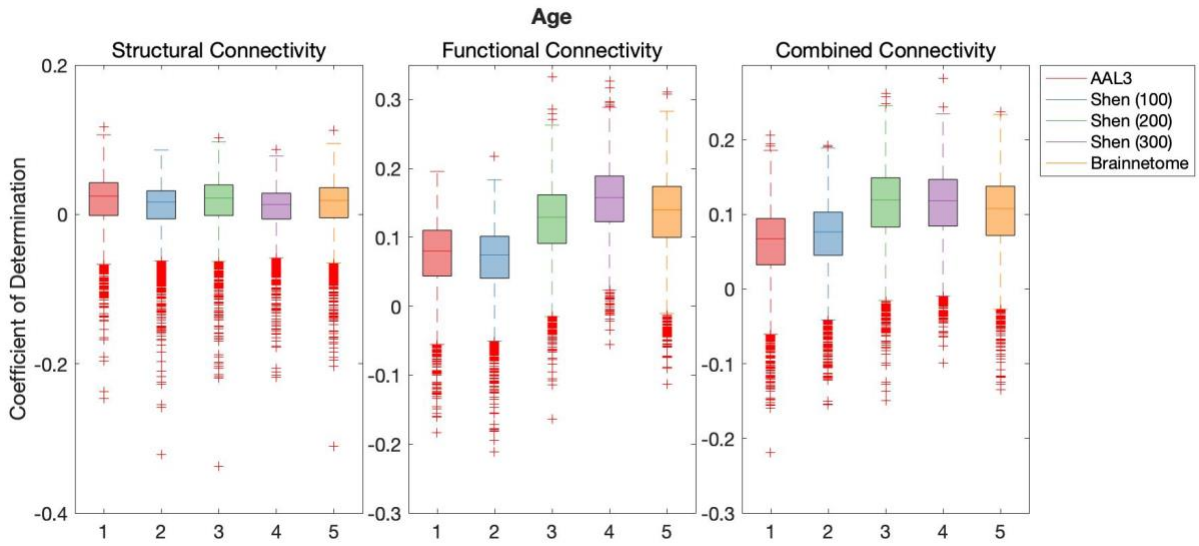

Figure 4 Results of BBC-CV of age, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate greater than chance prediction and unfilled boxes illustrate not greater than chance prediction.

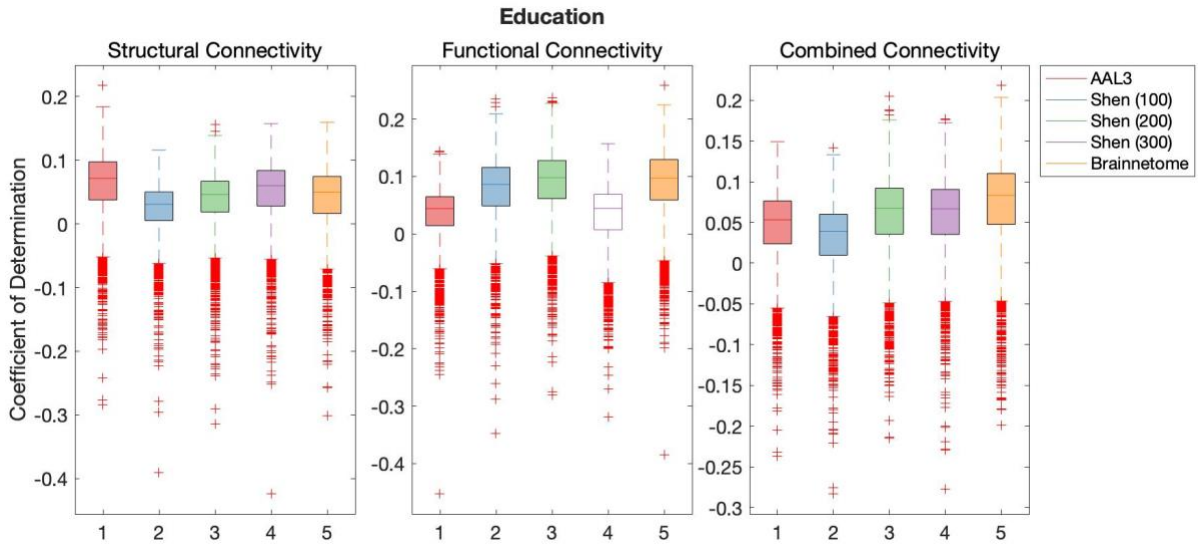

Figure 5 Results of BBC-CV of education, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate greater than chance prediction and unfilled boxes illustrate not greater than chance prediction.

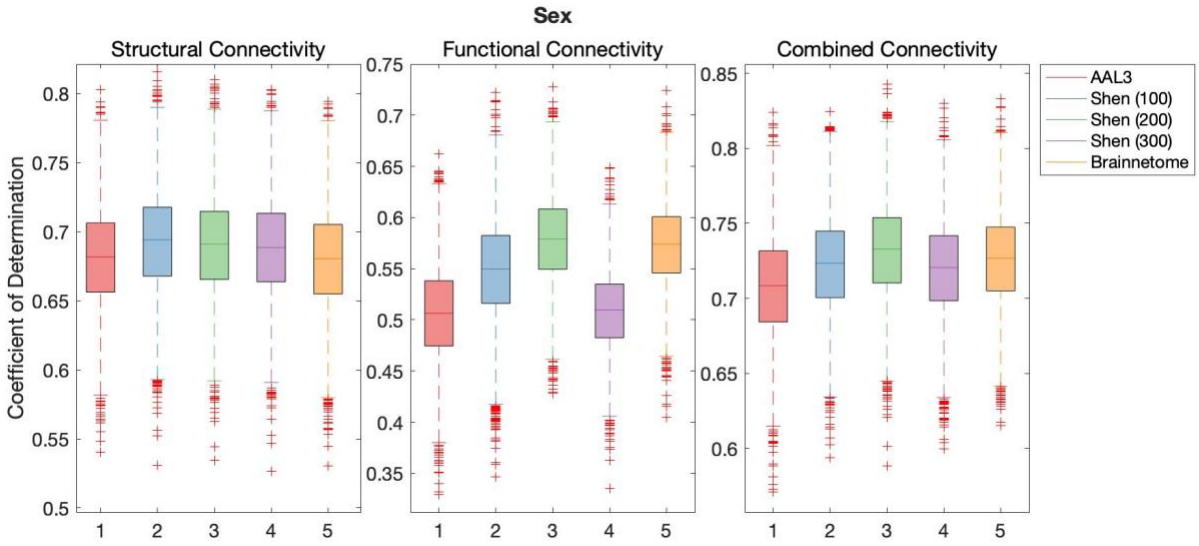

Figure 6 Results of BBC-CV of sex, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate greater than chance prediction and unfilled boxes illustrate not greater than chance prediction.

Table 1 Pairwise comparisons Z-scores of bootstrap-bias corrected out-of-sample predictions generated with each parcellation. Parcellations presented in rows were used as the reference, therefore positive values reflect that the parcellation presented in a given row performs better than the parcellation presented in a given column, whereas negative values indicate that the parcellation presented in a given row performs poorer than the parcellation presented in a given column. Comparisons were only performed for parcellations that produce better predictions than chance. All Z-values that are presented in bold have the p-value <.001.

Connectivity		**Age**				
		AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
Structural	AAL3	-	**14.07**	**3.76**	**18.47**	**9.27**
	Shen (100)	**-14.07**	-	**-10.60**	**4.66**	**-4.97**
	Shen (200)	**-3.76**	**10.60**	-	**15.22**	**5.65**
	Shen (300)	**-18.47**	**-4.66**	**-15.22**	-	**-9.61**
	Brainnetome	**-9.27**	**4.97**	**-5.65**	**9.61**	-
Functional	AAL3	-	**6.76**	**-44.40**	**-65.95**	**-50.66**
	Shen (100)	**-6.76**	-	**-50.22**	**-70.37**	**-55.70**
	Shen (200)	**44.40**	**50.22**	-	**-28.36**	**-10.03**
	Shen (300)	**65.95**	**70.37**	**28.36**	-	**17.78**

	Brainnetome	**50.66**	**55.70**	**10.03**	**-17.78**	-
Combined	AAL3	-	**-10.93**	**-48.91**	**-50.07**	**-40.62**
	Shen (100)	**10.93**	-	**-41.68**	**-42.57**	**-32.30**
	Shen (200)	**48.91**	**41.68**	-	**0.62**	**11.47**
	Shen (300)	**50.07**	**42.57**	**-0.62**	-	**11.14**
	Brainnetome	**40.62**	**32.30**	**-11.47**	**-11.14**	-
				Education		
		AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
Structural	AAL3	-	**45.49**	**30.57**	**13.74**	**24.61**
	Shen (100)	**-45.49**	-	**-21.13**	**-35.74**	**-24.15**
	Shen (200)	**-30.57**	**21.13**	**-**	**-18.15**	**-5.43**
	Shen (300)	**-13.74**	**35.74**	**18.15**	-	**11.95**
	Brainnetome	**-24.61**	**24.15**	**5.43**	**-11.95**	-
Functional	AAL3	-	**-43.94**	**-52.78**	-	**-51.19**
	Shen (100)	**43.94**	-	**-11.91**	-	**-11.59**
	Shen (200)	**52.78**	**11.91**	-	-	0.02
	Shen (300)	-	-	-	-	-
	Brainnetome	**51.19**	**11.59**	-0.02	-	-
Combined	AAL3	-	**19.49**	**-17.51**	**-16.18**	**-32.87**
	Shen (100)	**-19.49**	-	**-34.74**	**-33.69**	**-46.51**
	Shen (200)	**17.51**	**34.74**	-	1.49	**-17.27**
	Shen (300)	**16.18**	**33.69**	-1.49	-	**-18.70**
	Brainnetome	**32.87**	**46.51**	**17.27**	**18.70**	-
				Sex		
		AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
Structural	AAL3	-	**-15.09**	**-11.97**	**-9.36**	1.91
	Shen (100)	**15.09**	-	**3.45**	**5.98**	**16.85**
	Shen (200)	**11.97**	**-3.45**	-	**2.60**	**13.80**

	Shen (300)	**9.36**	**-5.98**	**-2.60**	-	**11.20**
	Brainnetome	-1.91	**-16.85**	**-13.80**	**-11.20**	-
Functional	AAL3	-	**-40.03**	**-63.69**	**-3.16**	**-61.65**
	Shen (100)	**40.03**	-	**-30.34**	**41.02**	**-25.60**
	Shen (200)	**63.69**	**30.34**	-	**66.43**	**6.52**
	Shen (300)	**3.16**	**-41.02**	**-66.43**	-	**-64.62**
	Brainnetome	**61.65**	**25.60**	**-6.52**	**64.62**	-
Combined	AAL3	-	**-21.11**	**-33.98**	**-17.36**	**-25.74**
	Shen (100)	**21.11**	-	**-14.08**	**4.12**	**-4.77**
	Shen (200)	**33.98**	**14.08**	-	**18.19**	**9.53**
	Shen (300)	**17.36**	**-4.12**	**-18.19**	-	**-8.96**
	Brainnetome	**25.74**	**4.77**	**-9.53**	**8.96**	-

Cognition

AIC-based model comparison

Figures 7-11 demonstrate AIC model evidence for models of cognition composed with SC, FC and CC defined with the 5 parcellation schemes.

When Executive Function was modelled, Shen 100 parcellation consistently outperformed other parcellation schemes. For SC models the difference relative to other parcellation schemes was positive (ΔAIC vs Shen 200 = -8.16), whereas the difference for FC and CC models relative to other parcellation schemes was weak (ΔAIC<= -3).

When Self-regulation was modelled, SC defined in Shen 100 parcellation outperformed alternative parcellation schemes (ΔAIC = -10.21). For FC, Shen 200 parcellation outperformed alternative parcellation schemes (ΔAIC = -6.13). For CC, Self-regulation could not be modelled with information obtained from Brainnetome parcellation in this dataset (all regression beta values were reduced to 0). Differences in model evidence for different parcellation schemes in CC were weak (ΔAIC < 3).

When Language was modelled, there was no notable difference between parcellation schemes used to define either SC or FC (ΔAIC < 3). However, there was some benefit of defining CC using Shen 100 parcellation relative to Shen 200 parcellation (ΔAIC = -3.09).

Encoding could not be modelled using AAL3 in SC or FC. In addition, there was no notable benefit of using any specific parcellation scheme in either SC or FC. However, for CC, AAL3 parcellation outperformed all alternative parcellation schemes (ΔAIC = -3.3).

Finally, Sequence Processing was most effectively modelled using Shen 100 parcellation in all modalities. However, this difference was barely worth a mention for FC and CC models (ΔAIC = -3.3).

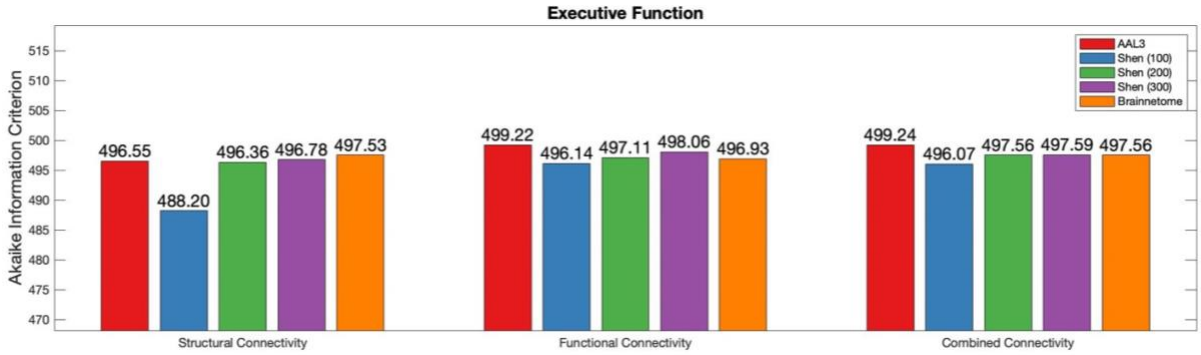

Figure 7 AIC model evidence for SC, FC and CC models of Executive Function, constructed with each parcellation scheme. Models with lower AIC values are favoured.

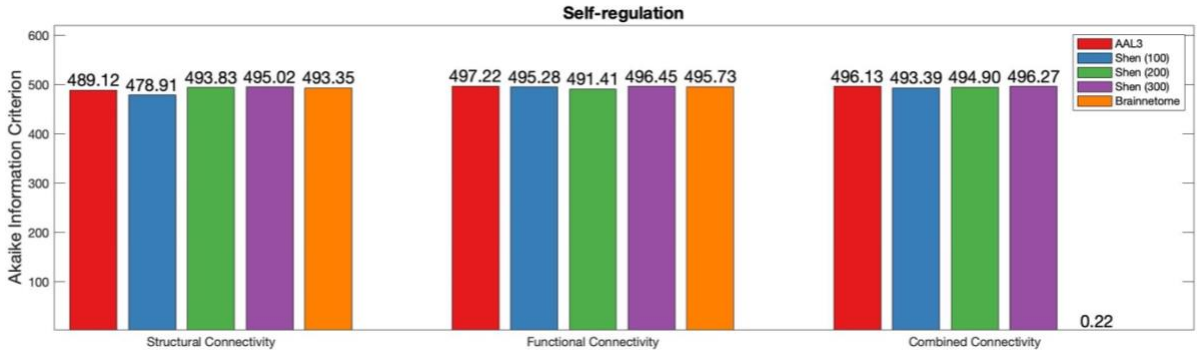

Figure 8 AIC model evidence for SC, FC and CC models of Self-regulation, constructed with each parcellation scheme. Models with lower AIC values are favoured.

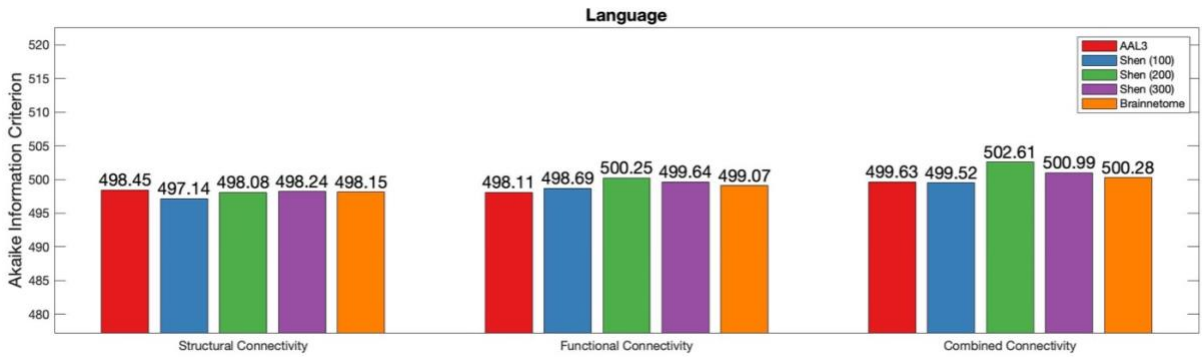

Figure 9 AIC model evidence for SC, FC and CC models of Language, constructed with each parcellation scheme. Models with lower AIC values are favoured.

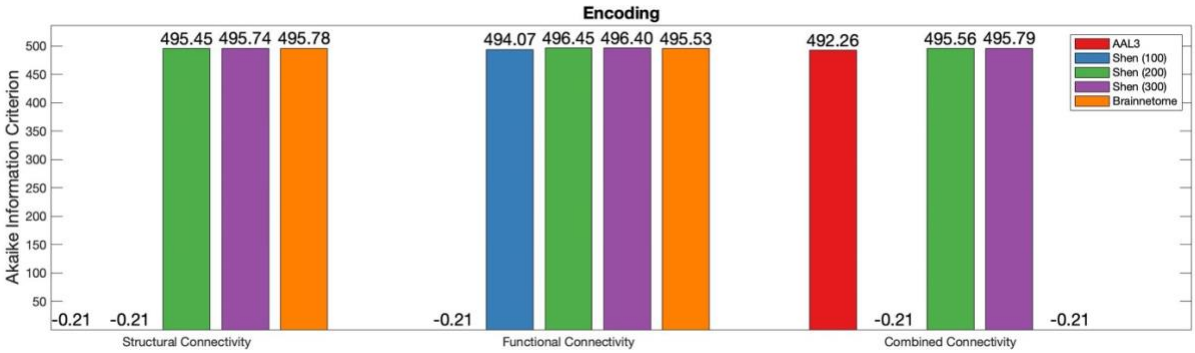

Figure 10 AIC model evidence for SC, FC and CC models of Encoding, constructed with each parcellation scheme. Models with lower AIC values are favoured.

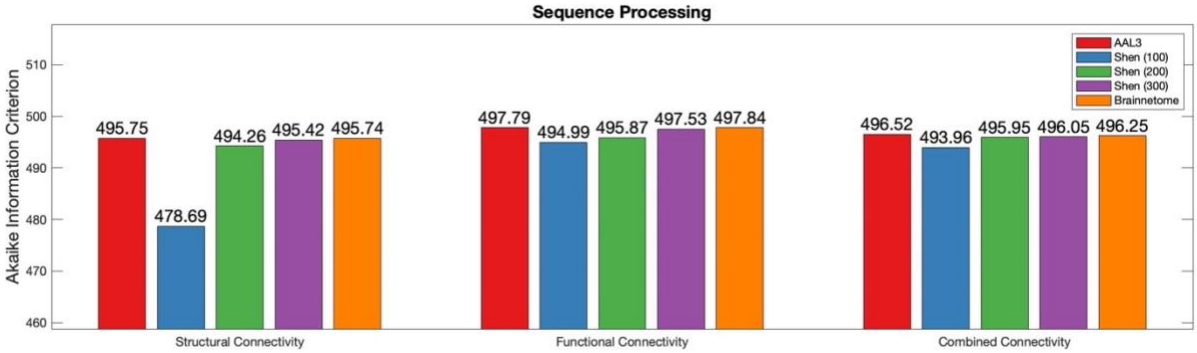

Figure 11 AIC model evidence for SC, FC and CC models of Sequence Processing, constructed with each parcellation scheme. Models with lower AIC values are favoured.

Cross-validation-based model comparison

Results of the cross-validation of models of cognition are presented in Figures 12-17 and pairwise comparisons of the cross-validation errors are presented in Table 2. The probability of pairwise comparisons was adjusted with False Discovery Rate (Benjamini & Hochberg, 1995). All following comparisons have passed the significance threshold of 0.0184.

Analysis has demonstrated that Shen 100 parcellation consistently produced out-of-sample predictions of cognition that explained more variance than chance. This was observed across structural, functional and combined connectivity. However, it rarely outperformed other parcellations at producing effective predictions of cognition (vs. Shen 200 Self-regulation structural connectivity 5.46, Language vs Brainnetome structural connectivity 3.39, functional connectivity 35.59 and combined connectivity 26.16, Language vs AAL3 with combined connectivity 42.56 and Language vs Shen 300 with combined connectivity 15.58). Next, AAL3 consistently produced the most effective predictions of Executive Function utilising functional connectivity (vs Shen 100 Z = 52.55, vs Shen 200 Z = 39.15, vs Brainnetome Z = 51.10) and combined connectivity (vs Shen 100 Z = 52.73, vs Shen 200 Z = 32.49, vs Brainnetome Z = 38.96). Similarly, Brainnetome consistently produced the most effective predictions of Self-regulation utilising structural connectivity (vs AAL 3 Z =8.56, Shen 100 Z = 9.21, vs Shen 200 Z = 16.65, vs Shen 300 Z = 14.65). Shen 200

parcellation consistently produced the most effective predictions of Language utilising functional connectivity (Shen 100 Z = 12.74, vs Brainnetome Z = 45.90). Shen 300 parcellation consistently produced the most effective predictions of Encoding utilising structural connectivity (Shen 100 Z = 9.45, vs Shen 200 Z = 12.19, vs Brainnetome Z = 10.91). There was no other parcellation scheme that consistently outperformed other parcellation schemes for any given type of connectivity or any cognitive construct.

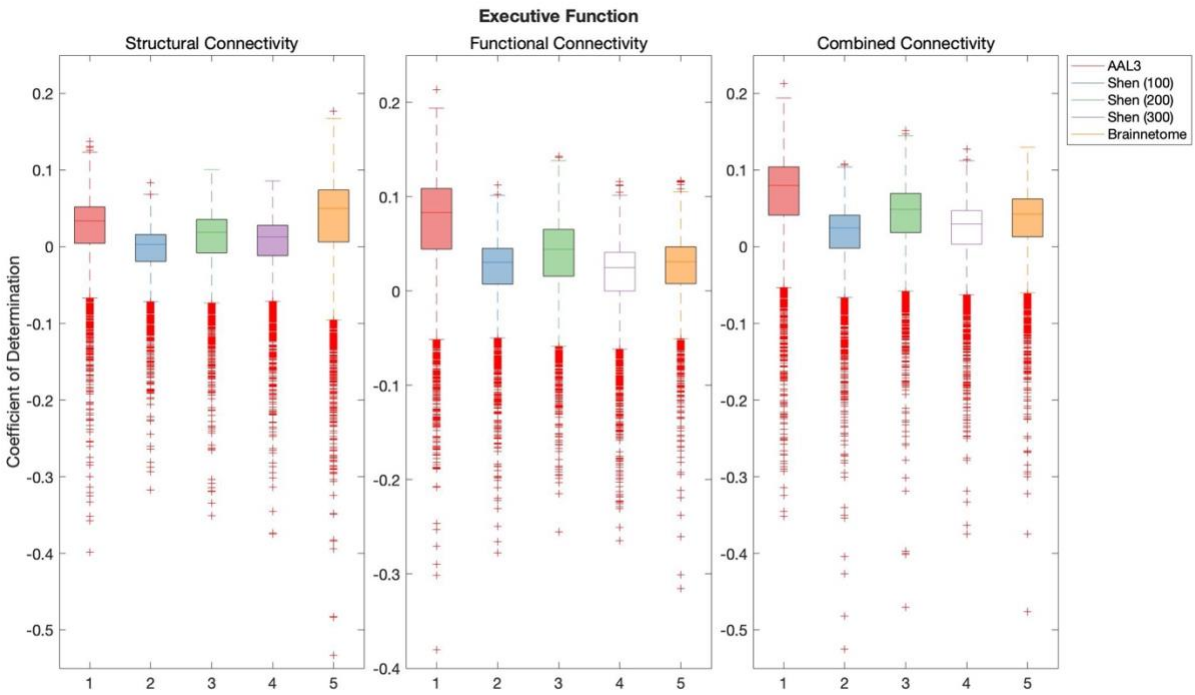

Figure 12 Results of BBC-CV of Executive Function, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate greater than chance prediction and unfilled boxes illustrate not greater than chance prediction.

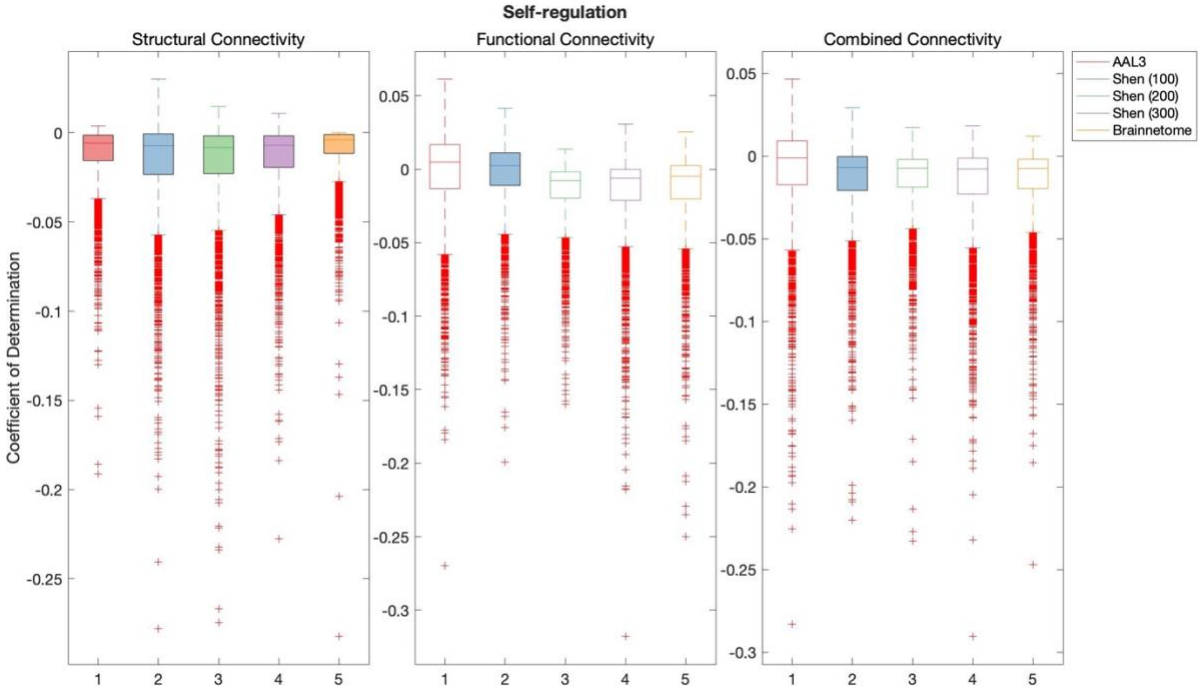

Figure 13 Results of BBC-CV of Self-regulation, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate greater than chance prediction and unfilled boxes illustrate not greater than chance prediction.

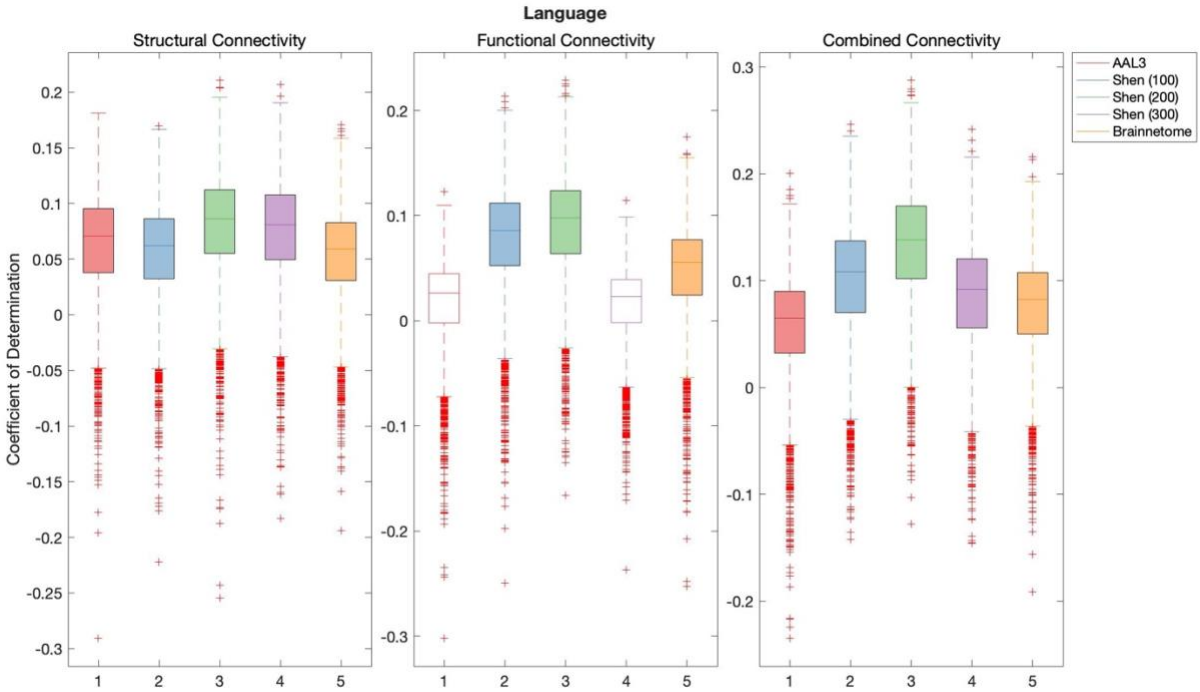

Figure 14 Results of BBC-CV of Language, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate greater than chance prediction and unfilled boxes illustrate not greater than chance prediction.

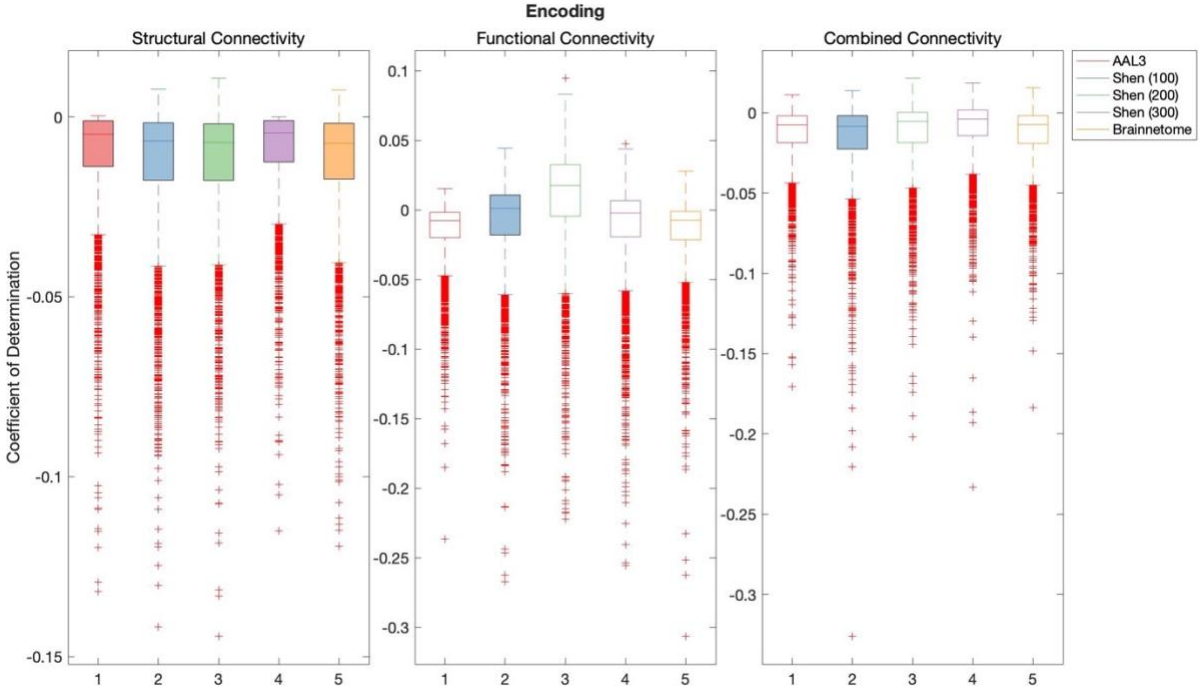

Figure 15 Results of BBC-CV of Encoding, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate greater than chance prediction and unfilled boxes illustrate not greater than chance prediction.

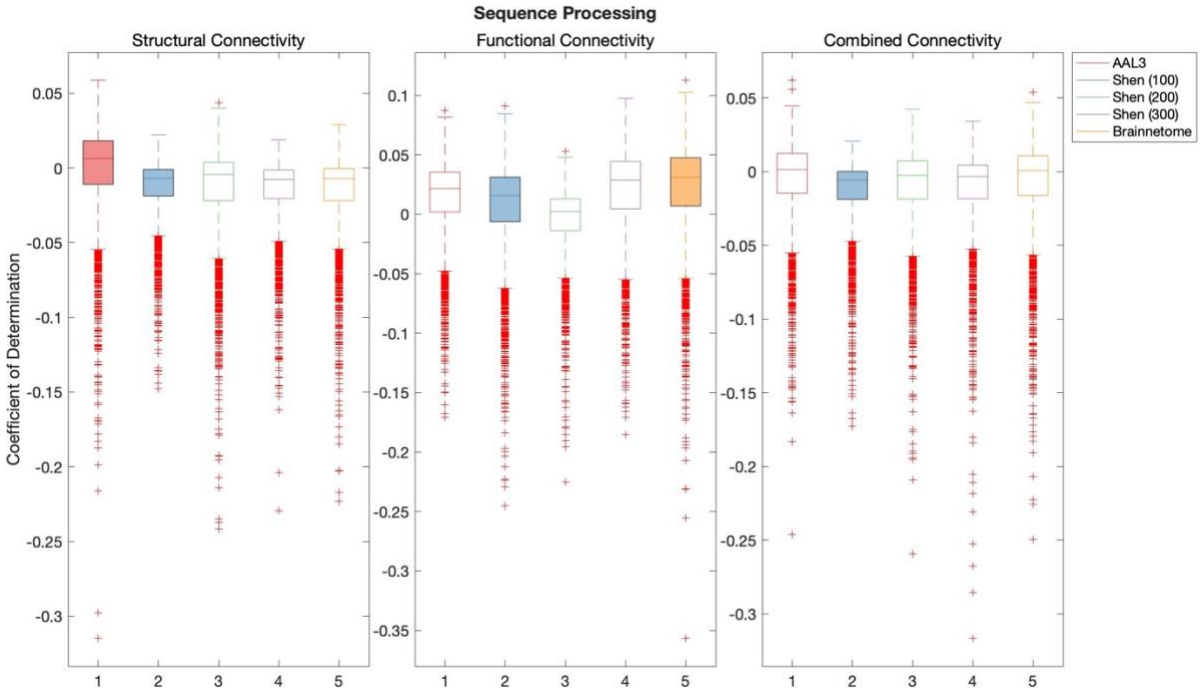

Figure 16 Results of BBC-CV of Sequence Processing, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate greater than chance prediction and unfilled boxes illustrate not greater than chance prediction.

Table 2 Pairwise comparisons Z-scores of bootstrap-bias corrected out-of-sample predictions generated with each parcellation. Parcellations presented in rows were used as the reference, therefore positive values reflect that the parcellation presented in a given row performs better than the parcellation presented in a given column, whereas negative values indicate that the parcellation presented in a given row performs poorer than the parcellation presented in a given column. Comparisons were only performed for parcellations that produce better predictions than chance. All Z-values that are presented in bold have the p-value <.001.

		Executive Function				
Connectivity		AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
Structural	AAL3	-	**43.85**	**21.77**	**31.34**	**-19.14**
	Shen (100)	**-43.85**	-	**-27.88**	**-18.63**	**-46.45**
	Shen (200)	**-21.77**	**27.88**	-	**11.09**	**-33.76**
	Shen (300)	**-31.34**	**18.63**	**-11.09**	-	**-39.49**
	Brainnetome	**19.14**	**46.45**	**33.76**	**39.49**	-
Functional	AAL3	-	**52.55**	**39.15**	-	**51.10**
	Shen (100)	**-52.55**	-	**-22.66**	-	-2.36
	Shen (200)	**-39.15**	**22.66**	-	-	**20.20**
	Shen (300)	-	-	-	-	-
	Brainnetome	**-51.10**	2.36	**-20.20**	-	-
Combined	AAL3	-	**52.73**	**32.49**	-	**38.96**
	Shen (100)	**-52.73**	-	**-33.88**	-	**-26.44**
	Shen (200)	**-32.49**	**33.88**	-	-	**8.85**
	Shen (300)	-	-	-	-	-
	Brainnetome	**-38.96**	**26.44**	**-8.85**	-	-
		Self-regulation				
		AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
Structural	AAL3	-	2.28	**8.78**	**6.21**	**-8.56**
	Shen (100)	-2.28	-	**5.46**	**2.86**	**-9.21**
	Shen (200)	**-8.78**	**-5.46**	-	**-2.95**	**-16.65**

	Shen (300)	**-6.21**	**-2.86**	**2.95**	-	**-14.65**
	Brainnetome	**8.56**	**9.21**	**16.65**	**14.65**	-
				Language		
		AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
Structural	AAL3	-	**9.33**	**-19.47**	**-13.50**	**12.71**
	Shen (100)	**-9.33**	-	**-28.35**	**-22.50**	**3.39**
	Shen (200)	**19.47**	**28.35**	-	**5.93**	**31.57**
	Shen (300)	**13.50**	**22.50**	**-5.93**	-	**25.76**
	Brainnetome	**-12.71**	**-3.39**	**-31.57**	**-25.76**	-
Functional	AAL3	-	**-**	**-**	-	**-**
	Shen (100)	**-**	-	**-12.74**	-	**35.59**
	Shen (200)	**-**	**12.74**	-	-	**45.90**
	Shen (300)	-	-	-	-	-
	Brainnetome	**-**	**-35.59**	**-45.90**	-	-
Combined	AAL3	-	**-42.56**	**-63.98**	**-29.41**	**-20.46**
	Shen (100)	**42.56**	-	**-30.72**	**15.58**	**26.16**
	Shen (200)	**63.98**	**30.72**	-	**43.99**	**52.63**
	Shen (300)	**29.41**	**-15.58**	**-43.99**	-	**10.74**
	Brainnetome	**20.46**	**-26.16**	**-52.63**	**-10.74**	-
				Encoding		
		AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
Structural	AAL3	-	**7.24**	**9.79**	-1.98	**8.57**
	Shen (100)	**-7.24**	-	1.98	**-9.45**	0.97
	Shen (200)	**-9.79**	-1.98	-	**-12.10**	-0.87
	Shen (300)	1.98	**9.45**	**12.10**	-	**10.91**
	Brainnetome	**-8.57**	-0.97	0.87	**-10.91**	-
				Sequence Processing		
		AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
Structural	AAL3	-	**38.51**	-	-	-

	Shen (100)	**-38.51**	-	-	-	-
	Shen (200)	-	-	-	-	-
	Shen (300)	-	-	-	-	-
	Brainnetome	-	-	-	-	-
Functional	AAL3	-	-	-	-	-
	Shen (100)	-	-	-	-	**-25.86**
	Shen (200)	-	-	-	-	-
	Shen (300)	-	-	-	-	-
	Brainnetome	-	**25.86**	-	-	-

Network organisation

Figures 17 and 18 present the distribution of scores obtained for each participant's global network measures across parcellations. The following Tables 3 and 4 present pair-wise comparisons made to assess differences in global graph theory measures across parcellations. It was generally found that parcellation schemes significantly impacted network organisation, as measured by all 6 graph theory measures. However, the direction of this impact was not necessarily shared across SC and FC. For example, the modularity statistic was highest when AAL3 was used to define SC and it was lowest when AAL3 was used to define FC. Correlation analysis between each global organisation measure and absolute model residuals obtained for models of demographics and cognition has not revealed any significant associations following correction for multiple comparisons.

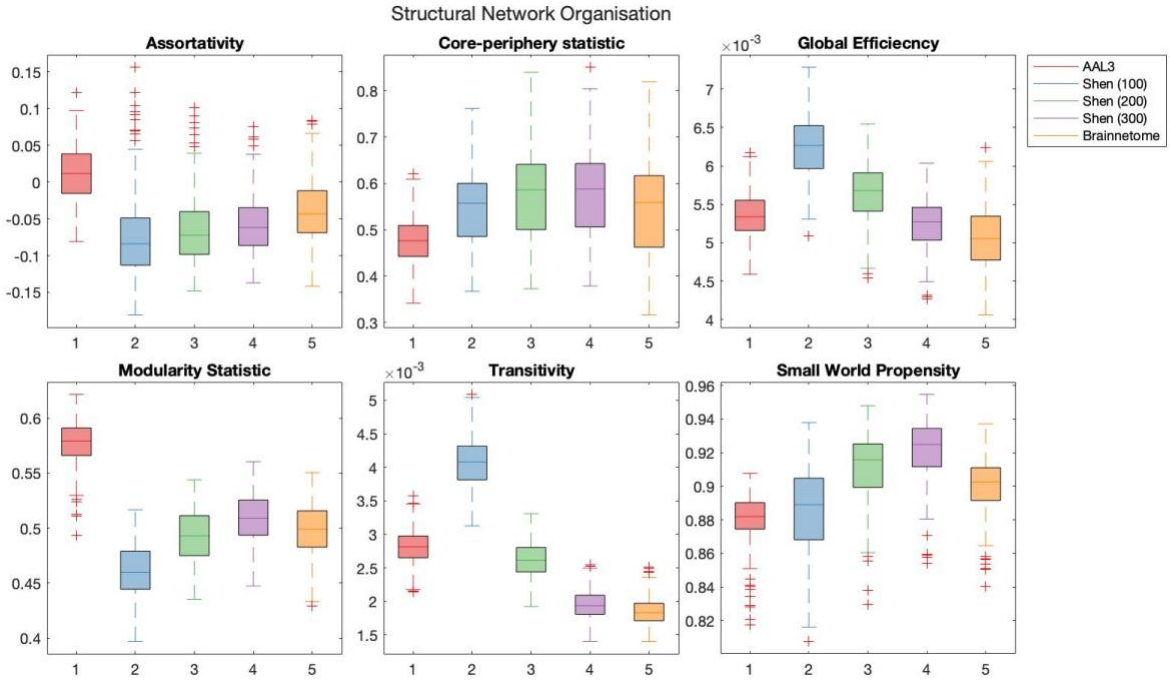

Figure 17 Graph theory measures of the global network organisation of SC.

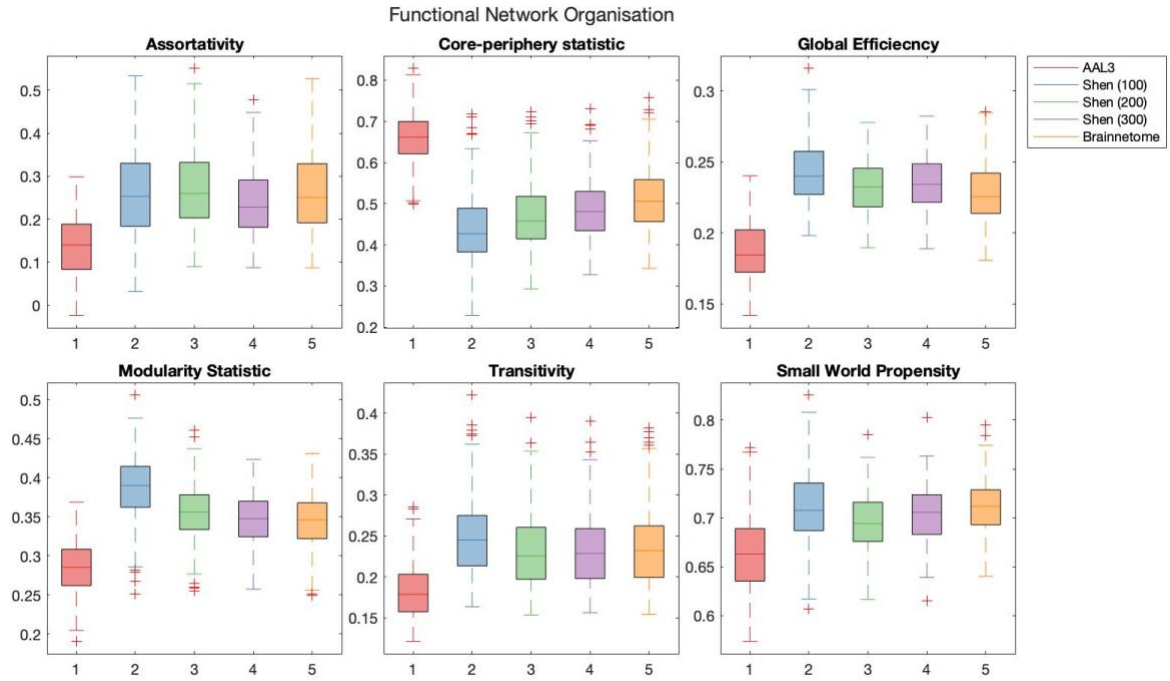

Figure 18 Graph theory measures of the global network organisation of FC.

Table 3 Pairwise comparison t-statistics of structural connectivity yielded by each network parcellation scheme. Parcellations presented in rows were used as the reference, therefore positive values reflect that the parcellation presented in a given row performs better than the parcellation presented in a given column, whereas negative values indicate that the parcellation presented in a given row performs poorer than the parcellation presented in a given

column. Comparisons were only performed for parcellations that produce better predictions than chance. All t-scores presented in bold have passed the FDR-adjusted critical p-value =<.0275. 0.0125

			Assortativity		
	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**25.16**	**24.66**	**24.26**	**19.27**
Shen (100)	**-25.16**	-	**-6.53**	**-8.27**	**-14.84**
Shen (200)	**-24.66**	**6.53**	-	**-6.94**	**-13.65**
Shen (300)	**-24.26**	**8.27**	**6.94**	-	**-10.22**
Brainnetome	**-19.27**	**14.84**	**13.65**	**10.22**	-
	Coreness Statistic				
	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**-23.34**	**-28.32**	**-29.75**	**-19.58**
Shen (100)	**23.34**	-	**-34.67**	**-34.91**	-0.75
Shen (200)	**28.32**	**34.67**	-	**-6.91**	**17.61**
Shen (300)	**29.75**	**34.91**	**6.91**	-	**21.08**
Brainnetome	**19.58**	0.75	**-17.61**	**-21.08**	-
	Global Efficiency				
	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**-51.98**	**-17.69**	**5.48**	**20.23**
Shen (100)	**51.98**	-	**82.15**	**108.86**	**71.26**
Shen (200)	**17.69**	**-82.15**	-	**93.32**	**32.85**
Shen (300)	**-5.48**	**-108.86**	**-93.32**	-	**10.62**
Brainnetome	**-20.23**	**-71.26**	**-32.85**	**-10.62**	-
	Modularity statistic				
	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**63.76**	**47.31**	**39.43**	**43.92**
Shen (100)	**-63.76**	-	**-32.99**	**-47.50**	**-29.80**
Shen (200)	**-47.31**	**32.99**	-	**-19.69**	**-5.07**
Shen (300)	**-39.43**	**47.50**	**19.69**	-	**9.32**
Brainnetome	**-43.92**	**29.80**	**5.07**	**-9.32**	-

	Transitivity				
	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**-68.75**	**11.51**	**61.78**	**75.16**
Shen (100)	**68.75**	-	**202.09**	**208.09**	**121.69**
Shen (200)	**-11.51**	**-202.09**	-	**183.45**	**56.06**
Shen (300)	**-61.78**	**-208.09**	**-183.45**	-	**7.82**
Brainnetome	**-75.16**	**-121.69**	**-56.06**	**-7.82**	-
	Small World Propensity				
	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**-3.98**	**-29.09**	**-43.74**	**-19.21**
Shen (100)	**3.98**	-	**-31.98**	**-37.98**	**-11.32**
Shen (200)	**29.09**	**31.98**	-	**-22.83**	**12.10**
Shen (300)	**43.74**	**37.98**	**22.83**	-	**26.27**
Brainnetome	**19.21**	**11.32**	**-12.10**	**-26.27**	-

Table 4 Pairwise comparison t-statistics of functional connectivity yielded by each network parcellation scheme. Parcellations presented in rows were used as the reference, therefore positive values reflect that the parcellation presented in a given row performs better than the parcellation presented in a given column, whereas negative values indicate that the parcellation presented in a given row performs poorer than the parcellation presented in a given column. Comparisons were only performed for parcellations that produce better predictions than chance. All t-scores presented in bold have passed the FDR-adjusted critical p-value =<.0275.

	Assortativity				
	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**-20.88**	**-21.62**	**-17.77**	**-21.62**
Shen (100)	**20.88**	-	**-2.63**	**6.39**	-0.09
Shen (200)	**21.62**	**2.63**	-	**13.90**	**3.00**
Shen (300)	**17.77**	**-6.39**	**-13.90**	-	**-8.80**
Brainnetome	**21.62**	0.09	**-3.00**	**8.80**	-
	Coreness Statistic				

	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**68.37**	**60.28**	**59.06**	**52.11**
Shen (100)	**-68.37**	-	**-18.33**	**-21.85**	**-32.32**
Shen (200)	**-60.28**	**18.33**	-	**-12.64**	**-27.03**
Shen (300)	**-59.06**	**21.85**	**12.64**	-	**-17.08**
Brainnetome	**-52.11**	**32.32**	**27.03**	**17.08**	-

Global Efficiency

	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**-84.70**	**-75.97**	**-76.49**	**-70.77**
Shen (100)	**84.70**	-	**22.21**	**14.90**	**31.20**
Shen (200)	**75.97**	**-22.21**	-	**-15.42**	**10.37**
Shen (300)	**76.49**	**-14.90**	**15.42**	-	**17.25**
Brainnetome	**70.77**	**-31.20**	**-10.37**	**-17.25**	-

Modularity statistic

	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**-56.89**	**-42.09**	**-37.50**	**-37.11**
Shen (100)	**56.89**	-	**21.18**	**23.92**	**27.29**
Shen (200)	**42.09**	**-21.18**	-	**7.55**	**9.30**
Shen (300)	**37.50**	**-23.92**	**-7.55**	-	**2.52**
Brainnetome	**37.11**	**-27.29**	**-9.30**	**-2.52**	-

Transitivity

	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**-51.56**	**-41.06**	**-42.83**	**-40.50**
Shen (100)	**51.56**	-	**22.37**	**19.87**	**15.01**
Shen (200)	**41.06**	**-22.37**	-	-0.46	**-7.33**
Shen (300)	**42.83**	**-19.87**	0.46	-	**-6.14**
Brainnetome	**40.50**	**-15.01**	**7.33**	**6.14**	-

Small World Propensity

	AAL3	Shen (100)	Shen (200)	Shen (300)	Brainnetome
AAL3	-	**-18.00**	**-15.90**	**-20.78**	**-24.86**

Shen (100)	**18.00**	-	**8.40**	**3.94**	0.69
Shen (200)	**15.90**	**-8.40**	-	**-8.88**	**-12.39**
Shen (300)	**20.78**	**-3.94**	**8.88**	-	**-5.67**
Brainnetome	**24.86**	-0.69	**12.39**	**5.67**	-

DISCUSSION

Various brain parcellation schemes are defined with different neuroimaging data and node-generation algorithms. Consequently, parcellation schemes influence the number, size and shape of nodes. Parcellation schemes have been demonstrated to largely impact the predictive power of connectivity-based models of cognition (Dhamala et al., 2021; Mellema et al., 2020; Ota et al., 2014; Pervaiz et al., 2020). However, to date, it remained unclear if the method of parcellation was systematically related to the efficiency of predictive models. To address this ambiguity, the present research has defined SC, FC and CC with 5 parcellation schemes and constructed predictive models of demographics and cognition. When AIC-model evidence was assessed, there was no notable benefit to choosing a specific parcellation to model sample demographics. However, there was some trend towards the benefit of using the low-resolution functional parcellation (Shen 100) when cognition was modelled. When cross-validation results were assessed, most parcellations produced more effective predictions of sample demographics than chance. The only model producing validation-sample predictions related to chance was the FC model of education level produced with high-resolution functional connectivity (Shen 300). In addition, it was found that models of demographics were most effective using FC defined with functional parcellations. There was no consistent benefit to the use of low or high-resolution parcellations for this modality. When SC was used to model age and education, AAL3 parcellation produced the most effective unseen-sample predictions. In contrast, models of sex benefited in SC modality from low-resolution functional parcellation. When unseen-sample predictions of cognition were assessed, it was found that no parcellation scheme would consistently produce less prediction error than all other schemes, even when the modality of connectivity and parcellation corresponded. However, unseen predictions of cognition made using connectivity defined with the Shen 100 parcellation scheme were consistently explaining more variance in the unseen sample than chance. This was the only parcellation scheme that achieved such consistency in performance.

Prior research has demonstrated that effective modelling of cognition to some extent depends on parcellation. For example, Dhamala et al. (2021) reported that SC defined with low-resolution FreeSurfer parcellation succeeded in producing predictions of crystallized intelligence in an unseen sample. However, the same connectivity failed to produce effective predictions when in-house high-resolution parcellation (CoCo 439) was used instead. In contrast, FC tended to benefit at predicting cognitive abilities when CoCo 439 parcellation was used. Their CoCo 439 parcellation was defined from functional data

suggesting a benefit of using a parcellation that corresponds to the modality of connectivity. However, in the present report, when individual cognitive domains were analysed such benefits were lost. For example, we found that Executive Function was most effectively modelled with FC when FC was defined with AAL3 parcellation – a parcellation defined with neural anatomy. In addition, similarly to the report from Dhamala et al. (2021), this work could not identify a single superior parcellation. Such inconsistency of findings warrants caution when interpreting results obtained with predictive modelling. Authors must consider that any superiority of a given model defined with a given set of data may be related to the parcellation scheme used.

In another seminal report, Pervaiz et al. (2020) demonstrated that the most accurate predictions of fluid cognition were produced with models defined with high-resolution parcels generated by spatial Independent Component Analysis. With this algorithm, parcels can overlap and are not necessarily contiguous. This contrasts with publicly available parcellations used in the present work, which were defined to identify non-overlapping, contiguous parcels. Consequently, the parcellations offered with fewer constraints may prove to produce more accurate parcels with coherent signals (i.e. avoid partial voluming of signals). For example, recent work has demonstrated graded changes in connectivity patterns across cortex (Bajada et al., 2017; Cloutman et al., 2020). This may explain why the results presented by Pervaiz et al. (2020) identified high resolution to benefit modelling of fluid cognition, while our results could not identify this pattern. However, Pervaiz and colleagues (2020) have not assessed the predictive value of various parcellation schemes in modelling done with different parcellations with SC or CC. The present work adds to the field the finding that SC and CC do not benefit in predictive power from modality-corresponding parcellation or a specific resolution of parcellation.

One consistent finding across models of demographics and cognition was that the low-resolution functional parcellation (Shen 100) always produced predictions that explained more variation than chance. This was observed regardless of whether SC, FC or CC was used to predict demographics and cognition. One explanation of this finding is that higher-resolution parcellations may be more prone to partial voluming of signals due to individual differences in neural anatomy and function. This would result in models that are more specific to the training sample and decreased generalisability skills. However, where higher-resolution parcellation schemes succeeded at modelling demographics and cognition, they always proved to produce more generalisable models. This highlights that both low-resolution parcellations may offer some insights into neural substrates of cognition and high-resolution further complement this information.

Some whole sample models of Self-regulation and Encoding have been unable to effectively model cognition with the whole sample. This occurs due to elastic net regularisation of regression's beta coefficients that can select a severe reduction of beta coefficients to zero. The severity of regularisation has been defined by tuning alpha and lambda values with cross-validation. The fact that alpha and lambda favoured L1 penalty,

resulting in 0-beta coefficients, likely reflects the pattern that simpler models are needed to achieve good generalisability to validation samples when a smaller number of training data is available during the modelling of Self-regulation and Encoding. An increase in training sample size may benefit these specific models and remedy the problem. Alternatively, ridge regression could be used, which does not reduce any beta coefficients to zero. However, such a change in the procedure, without increasing the training sample size, would likely result in models that have poor generalisability.

One surprising finding of this work was the finding that although parcellation schemes have impacted the global architecture of brain networks, this had no association with model residuals for the models. In other words, once variance in cognition has been accounted for by connectivity, there was little to no relationship between cognition and network organisation. Although there were some trends towards associations. This suggests that global brain organisation measures reflect changes in the organisation across parcellation schemes but this change alone cannot account for the reason why specific parcellation schemes excel at modelling cognition. At the moment, partial voluming of signals presents a much stronger explanation of why specific parcellations succeed at modelling cognition and others fail. In the future, it will be possible to implement in-house parcellations that allow for overlapping parcels and manipulate the resolution of parcellation (e.g. spatial Independent Component Analysis, spectral clustering and spectral-reordering algorithm), to generate different parcellations and investigate further if organisational properties revealed by specific resolution of parcellation relate to demographics and cognition.

One key limitation sets the boundary of the interpretation of the present results. We have not considered alternative resolutions of parcellations defined with anatomical or a hybrid of anatomical and functional information. In particular, we have investigated 3 resolutions of functional parcellation but only a single resolution of anatomical and hybrid parcellation. It is possible that investigations of alternative resolutions in these corresponding modalities of connectivity would reveal a distinct benefit for predictive modelling with SC or CC. For example, it is possible that high- or low-resolution SC is particularly important for the prediction of cognitive domains and we failed to capture this because we only varied resolution for the functional parcellation scheme. A functional parcellation scheme is likely poorly suited to accurately identify parcel volumes in structural connectivity and it may have missed the benefit of various resolutions of structural parcellations.

Overall, this paper has investigated if the method of parcellation is systematically related to the efficiency of predictive models of cognition and parcellation. By investigating the generalisability of SC, FC and CC predictive models defined with 5 different parcellations, we have found that low-resolution parcellation (Shen 100) was consistently able to produce generalisable models of demographics and cognition. However, when alternative parcellations succeeded at producing generalisable models then these

parcellations outperformed the Shen 100 parcellation. It remained challenging to identify a superior parcellation scheme. As a result of this work, it is important that from here onwards authors provide a cautious interpretation of superior models of cognition based on connectivity, as specific parcellation schemes may unsystematically impact the results.

References

Akiki, T. J., & Abdallah, C. G. (2019). Determining the Hierarchical Architecture of the Human Brain Using Subject-Level Clustering of Functional Networks. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-55738-y

Alvarez, P., & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: a simple network model. *Proceedings of the National Academy of Sciences*, *91*(15), 7041-7045. https://doi.org/10.1073/pnas.91.15.7041

Amunts, K., & Zilles, K. (2015). Architectonic Mapping of the Human Brain beyond Brodmann. *Neuron*, *88*(6), 1086-1107. https://doi.org/10.1016/j.neuron.2015.12.001

Arslan, S., Ktena, S. I., Makropoulos, A., Robinson, E. C., Rueckert, D., & Parisot, S. (2018). Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex. *NeuroImage*, *170*, 5-30. https://doi.org/10.1016/j.neuroimage.2017.04.014

Ashourvan, A., Telesford, Q. K., Verstynen, T., Vettel, J. M., & Bassett, D. S. (2019). Multi-scale detection of hierarchical community architecture in structural and functional brain networks. *PLoS One*, *14*(5), e0215520. https://doi.org/10.1371/journal.pone.0215520

Bajada, C. J., Jackson, R. L., Haroon, H. A., Azadbakht, H., Parker, G. J. M., Lambon Ralph, M. A., & Cloutman, L. L. (2017). A graded tractographic parcellation of the temporal lobe. *NeuroImage*, *155*, 503-512. https://doi.org/10.1016/j.neuroimage.2017.04.016

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Brodmann, K. (1909). Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. *Leipzig: Barth*.

Cloutman, L. L., Lambon Ralph, M. A., Bajada, C. J., & Jackson, R. L. (2020). The Graded Change in Connectivity across the Ventromedial Prefrontal Cortex Reveals Distinct Subregions. *Cerebral Cortex*, *30*(1), 165-180. https://doi.org/10.1093/cercor/bhz079

Cohen, A. L., Fair, D. A., Dosenbach, N. U. F., Miezin, F. M., Dierker, D., Van Essen, D. C., Schlaggar, B. L., & Petersen, S. E. (2008). Defining functional areas in individual human brains using resting functional connectivity MRI. *NeuroImage*, *41*(1), 45-57. https://doi.org/10.1016/j.neuroimage.2008.01.066

de Kwaasteniet, B., Ruhe, E., Caan, M., Rive, M., Olabarriaga, S., Groefsema, M., Heesink, L., van Wingen, G., & Denys, D. (2013). Relation Between Structural and Functional Connectivity in Major Depressive Disorder. *Biological Psychiatry*, *74*(1), 40-47. https://doi.org/10.1016/j.biopsych.2012.12.024

Dhamala, E., Jamison, K. W., Jaywant, A., Dennis, S., & Kuceyeski, A. (2021). Distinct functional and structural connections predict crystallised and fluid cognition in healthy adults. *Human Brain Mapping*, *42*(10), 3102-3118. https://doi.org/10.1002/hbm.25420

Dickie, D. A., Shenkin, S. D., Anblagan, D., Lee, J., Blesa Cabez, M., Rodriguez, D., Boardman, J. P., Waldman, A., Job, D. E., & Wardlaw, J. M. (2017). Whole Brain Magnetic Resonance Image Atlases: A Systematic Review of Existing Atlases and Caveats for Use in Population Imaging. *Frontiers in Neuroinformatics*, *11*. https://doi.org/10.3389/fninf.2017.00001

Eickhoff, S. B., Yeo, B. T. T., & Genon, S. (2018). Imaging-based parcellations of the human brain. *Nature Reviews Neuroscience*, *19*(11), 672-686. https://doi.org/10.1038/s41583-018-0071-7

Farahani, F. V., Karwowski, W., & Lighthall, N. R. (2019). Application of Graph Theory for Identifying Connectivity Patterns in Human Brain Networks: A Systematic Review. *Frontiers in Neuroscience*, *13*. https://doi.org/10.3389/fnins.2019.00585

Fornito. (2010). Network scaling effects in graph analytic studies of human resting-state fMRI data. *Frontiers in Systems Neuroscience*. https://doi.org/10.3389/fnsys.2010.00022

Guye, M., Bettus, G., Bartolomei, F., & Cozzone, P. J. (2010). Graph theoretical analysis of structural and functional connectivity MRI in normal and pathological brain networks. *Magnetic Resonance Materials in Physics, Biology and Medicine*, *23*(5-6), 409-421. https://doi.org/10.1007/s10334-010-0205-z

Hahn, K., Myers, N., Prigarin, S., Rodenacker, K., Kurz, A., Förstl, H., Zimmer, C., Wohlschläger, A. M., & Sorg, C. (2013). Selectively and progressively disrupted structural connectivity of functional brain networks in Alzheimer's disease — Revealed by a novel framework to analyze edge distributions of networks detecting disruptions with strong statistical evidence. *NeuroImage*, *81*, 96-109. https://doi.org/10.1016/j.neuroimage.2013.05.011

Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R., & Hagmann, P. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences*, *106*(6), 2035-2040. https://doi.org/10.1073/pnas.0811168106

Hoppenbrouwers, M., Vandermosten, M., & Boets, B. (2014). Autism as a disconnection syndrome: A qualitative and quantitative review of diffusion tensor imaging studies. *Research in Autism Spectrum Disorders*, *8*(4), 387-412. https://doi.org/10.1016/j.rasd.2013.12.018

Keller, T. A., Kana, R. K., & Just, M. A. (2007). A developmental study of the structural integrity of white matter in autism. *NeuroReport*, *18*(1), 23-27. https://doi.org/10.1097/01.wnr.0000239965.21685.99

Lawrence, R. M., Bridgeford, E. W., Myers, P. E., Arvapalli, G. C., Ramachandran, S. C., Pisner, D. A., Frank, P. F., Lemmer, A. D., Nikolaidis, A., & Vogelstein, J. T. (2021). Standardizing human brain parcellations. *Scientific Data*, *8*(1). https://doi.org/10.1038/s41597-021-00849-3

Litwińczuk, M. C., Trujillo-Barreto, N., Muhlert, N., Cloutman, L., & Woollams, A. (2022). Combination of structural and functional connectivity explains unique variation in specific domains of cognitive function. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2022.119531

Manca, R., Sharrack, B., Paling, D., Wilkinson, I. D., & Venneri, A. (2018). Brain connectivity and cognitive processing speed in multiple sclerosis: A systematic review. *Journal of the Neurological Sciences*, *388*, 115-127. https://doi.org/10.1016/j.jns.2018.03.003

Mellema, C. J., Treacher, A., Nguyen, K. P., & Montillo, A. (2020). *Architectural configurations, atlas granularity and functional connectivity with diagnostic value in*

Autism Spectrum Disorder 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI),

Muldoon, S. F., Bridgeford, E. W., & Bassett, D. S. (2016). Small-World Propensity and Weighted Brain Networks. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep22057

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

Onnela, J. P., Saramaki, J., Kertesz, J., & Kaski, K. (2005). Intensity and coherence of motifs in weighted complex networks. *Phys Rev E Stat Nonlin Soft Matter Phys*, *71*(6 Pt 2), 065103. https://doi.org/10.1103/PhysRevE.71.065103

Ota, K., Oishi, N., Ito, K., & Fukuyama, H. (2014). A comparison of three brain atlases for MCI prediction. *Journal of Neuroscience Methods*, *221*, 139-150. https://doi.org/10.1016/j.jneumeth.2013.10.003

Park, H.-J., & Friston, K. (2013). Structural and Functional Brain Networks: From Connections to Cognition. *Science*, *342*(6158). https://doi.org/10.1126/science.1238411

Pervaiz, U., Vidaurre, D., Woolrich, M. W., & Smith, S. M. (2020). Optimising network modelling methods for fMRI. *NeuroImage*, *211*. https://doi.org/10.1016/j.neuroimage.2020.116604

Poldrack, R. A., Huckins, G., & Varoquaux, G. (2020). Establishment of Best Practices for Evidence for Prediction. *JAMA Psychiatry*, *77*(5). https://doi.org/10.1001/jamapsychiatry.2019.3671

Rolls, E. T., Huang, C.-C., Lin, C.-P., Feng, J., & Joliot, M. (2020). Automated anatomical labelling atlas 3. *NeuroImage*, *206*. https://doi.org/10.1016/j.neuroimage.2019.116189

Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J., & Joliot, M. (2020). Automated anatomical labelling atlas 3. *NeuroImage*, *206*, 116189. https://doi.org/10.1016/j.neuroimage.2019.116189

Salami, A., Rieckmann, A., Fischer, H., & Bäckman, L. (2014). A multivariate analysis of age-related differences in functional networks supporting conflict resolution. *NeuroImage*, *86*, 150-163. https://doi.org/10.1016/j.neuroimage.2013.08.002

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., & Greicius, M. D. (2007). Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. *Journal of Neuroscience*, *27*(9), 2349-2356. https://doi.org/10.1523/jneurosci.5587-06.2007

Shen, X., Tokoglu, F., Papademetris, X., & Constable, R. T. (2013). Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *NeuroImage*, *82*, 403-415. https://doi.org/10.1016/j.neuroimage.2013.05.081

Tsamardinos, I., Greasidou, E., & Borboudakis, G. (2018). Bootstrapping the out-of-sample predictions for efficient and accurate cross-validation. *Machine Learning*, *107*(12), 1895-1922. https://doi.org/10.1007/s10994-018-5714-4

van den Heuvel, M. P., & Fornito, A. (2014). Brain Networks in Schizophrenia. *Neuropsychology Review*, *24*(1), 32-48. https://doi.org/10.1007/s11065-014-9248-7

Wang, J., Wang, L., Zang, Y., Yang, H., Tang, H., Gong, Q., Chen, Z., Zhu, C., & He, Y. (2009). Parcellation-dependent small-world brain functional networks: A resting-state fMRI study. *Human Brain Mapping*, *30*(5), 1511-1523. https://doi.org/10.1002/hbm.20623

Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS One*, *8*(7), e68910. https://doi.org/10.1371/journal.pone.0068910

Yao, Z., Hu, B., Xie, Y., Moore, P., & Zheng, J. (2015). A review of structural and functional brain networks: small world and atlas. *Brain Informatics*, *2*(1), 45-52. https://doi.org/10.1007/s40708-015-0009-z

Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yücel, M., Pantelis, C., & Bullmore, E. T. (2010). Whole-brain anatomical networks: Does the choice of nodes matter? *NeuroImage*, *50*(3), 970-983. https://doi.org/10.1016/j.neuroimage.2009.12.027