Impact of brain parcellation on prediction performance 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 of a network whose connections will be studied. Brain connectivity has already been used in predictive modelling of cognition, but it remains unclear if the resolution of the parcellation used can systematically impact the predictive model performance. In this work, structural, functional and combined connectivity were each defined with 5 different parcellation schemes. The resolution and modality of the parcellation schemes were varied. Each connectivity defined with each parcellation was used 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 across all cognitive domains and demographics. In addition, although parcellation schemes impacted the global organisation of each connectivity type, this difference could not account for the out-of-sample prediction performance of the models. Taken together, these findings demonstrate that while high-resolution parcellations may be beneficial for modelling specific individual differences, partial voluming of signals produced 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, research investigates the white matter connections between regions (structural connectivity, SC) and the statistical associations between their activity (functional connectivity, FC)(Park & Friston, 2013). SC and FC have been related to healthy cognitive function throughout the lifespan (Salami et al., 2014), whereas their disruption 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 to define the brain regions (i.e. brain parcellation) used for connectivity analysis (Eickhoff et al., 2018; Yao et al., 2015). In the analysis of brain connectivity, the brain is idealised (modelled) as a set of interconnected nodes each representing a brain region that is

assumed to constitute a coherent unit of activity. The connectivity pattern between regions is then studied and often used to explain and predict cognitive function. Therefore, brain parcellation is a key step of any 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 based on different parcellation criteria. Consequently, each parcellation offers different node locations, sizes and shapes therefore leading to different representations of the activity units used for analysis. 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 highlight specific properties of the brain (Eickhoff et al., 2018). To illustrate, we may choose to parcellate the brain based on histological boundaries (Amunts & Zilles, 2015; Brodmann, 1909) or clustering of a group of similarly active voxels (Cohen et al., 2008). Such choices appear to impact consistency of research findings (Dhamala et al., 2021; Mellema et al., 2020; Ota et al., 2014; Pervaiz et al., 2020).

Despite the lack of an accepted standard, parcellation choice is an important decision in network analysis and it carries profound consequences for interpretation of results and comparison across different studies. It has been empirically demonstrated that parcellation can impact the properties of the estimated brain networks. For example, the parcellation used can impact 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, the ability to produce highly connected nodes (aka degree distribution) can also be affected by the specific parcellation used (Fornito, 2010; Wang et al., 2009; Zalesky et al., 2010). Similarly, the choice of parcellation resolution has also been demonstrated to impact the similarities between structural and functional connectivity (Akiki & Abdallah, 2019; Ashourvan et al., 2019; Honey et al., 2009).

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 parcellations may benefit the predictive power of connectivity models of cognition (Mellema et al., 2020; Pervaiz et al., 2020). The difference in predictive power of various brain parcellations may be related to the properties of the networks that become more prominent with specific parcellations. For example, increased small world characteristics of FC have been related to high intelligence (Hilger et al., 2017; Langer et al., 2012). Therefore, if one brain parcellation tends to generate networks with increased specific properties, then this parcellation may influence the accuracy of model predictions. However, it is also possible that difference in predictive power of specific models of cognition is related to whether specific parcellations generate

accurate connectivity profiles. For example, it is likely that a functionally-defined parcellation is less prone to partial voluming of functional signals as compared to structurally-defined parcellation. In contrast, a structurally-defined parcellation that considers anatomical boundaries of regions may be better suited to define SC. In support of this proposal, Dhamala et al. (2021) have found a trend that SC defined with structural parcellation yielded more effective models of crystallised cognition than SC defined with activation-based parcellation.

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 parcellations: the resting state activity-defined parcellation with either 93, 184 and 278 parcels (Shen et al., 2013), the structurally-defined parcellation (Rolls et al., 2020) and a hybrid parcellation that considers both structural and functional information (Xia et al., 2013). Our work then produced connectivity-based predictive models of cognition and used graph theory to assess the global differences in network properties that may explain model differences in model performance. Doing so, we tested the hypotheses that (1) SC-based models of demographics and cognition will be more effective when connectivity is defined with a structural parcellation (Rolls et al., 2020) than with a parcellation based on a functional activation (Shen et al., 2013), (2) FC-based models of demographics and cognition will be more effective when connectivity is defined with a parcellation based on brain function (Shen et al., 2013) than with a parcellation based on structure, (3) use of high-resolution parcellation will improve explanatory and predictive power of models of demographics and cognition regardless of connectivity type, (4) 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 same measures of cognition as our previous work (Litwińczuk et al., 2022). 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). Principal Component Analysis (PCA) with VARIMAX rotation was then used as a feature extraction method and applied to the behavioural dataset. 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

The Neuroimaging data were further processed following the same processing pipeline as in our previous work (Litwińczuk et al., 2022). This pipeline is summarised for completeness in the next two sub-sections. Both SC and FC were defined with 5 parcellation schemes that varied in their modality and resolution. For clarity, throughout this work the parcellation name will be assisted with the number of parcels:

1. Low-resolution resting-state functional parcellation composed of 93 parcels (Shen et al., 2013)

2. Structural parcellation known as Automated Anatomical Labelling (AAL3) composed of 166 parcels (Rolls et al., 2020)
3. Moderate-resolution resting-state functional parcellation composed of 184 parcels (Shen et al., 2013)
4. Hybrid structural and functional parcellation known as Brainnetome composed of 246 parcels (Xia et al., 2013).
5. High-resolution resting-state functional parcellation composed of 278 parcels (Shen et al., 2013)

Structural data and structural connectivity calculation

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 in each of the 5 parcellations.

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 (S_{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 ($S_{ij} \neq S_{ji}$), 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 ($S_{ij} + S_{ji}$), 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):,

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

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 $S_{ij} = N_i * 5000$ and $S_{ji} = N_j * 5000$, 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 rs-fMRI data was used 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 corrected, 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 5 parcellations. The average blood oxygenation level-dependent signal in each parcel was obtained and the pairwise (parcel-to-parcel) 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

Predictive models of 5 cognitive domains and 3 demographic variables were separately constructed using FC, SC or a combination of SC and FC (referred to as combined connectivity (CC)) as predictors. Models also differed in the brain parcellation scheme used to define the network nodes for connectivity calculations. This led to a total of 120 models to be analysed ((5 cognitive domains + 3 demographic characteristics) * 3 connectivity types * 5 parcellation schemes). Prior to model estimation, for each cognitive domain, the confounding effect of age, gender and education was regressed out of the response variable. Meanwhile, for each demographic, the remaining demographics were included in the main model as covariates of no interest. Then, all models were estimated using a Principal Component Regression (PCR) approach with elastic net regularisation of the

regression coefficients in latent space (EN-PCR). The overall estimation procedure consisted of the following steps:

1) PCA decomposition was used to orthogonalise the predictors' (connectivity) matrix
2) A regression model in latent (PCA) space was fitted with elastic net regularisation for the regression coefficients
3) Regression coefficients obtained in PCA space were projected back to the original connectome space and used to produce predicted responses.

To tune the Elastic Net regularisation hyper-parameters alpha and lambda and evaluate the out-of-sample model performance, a Bootstrap Bias Corrected Cross-Validation (BBC-CV) approach was used (Tsamardinos et al., 2018). In brief, the BBC-CV consisted in a repeated (50 times) 5-fold cross-validation with hyperparameter tuning (CVT), followed by a Bootstrap bias correction procedure (5000 bootstrap samples). The later Bootstrap step accounts for optimistic biases in the estimation of model performance introduced by using the same data for both hyperparameter tuning and model evaluation in the CVT step (Tsamardinos et al., 2018). The explained variation (coefficient of determination) between the predicted and observed responses was used as the statistic for both hyperparameter tuning and out-of-sample performance evaluation of the model (Poldrack et al., 2020). The resultant population of bootstrapped statistics was used to produce mean performance estimates of the EN-PCR learning strategy and corresponding confidence intervals. For a similar application of the BBC-CV method in the context of predictive modelling of cognition based on connectivity we refer the authors to our previous paper (Litwińczuk et al., 2022).

Finally, a permutation test was used to assess how likely it is to get the observed models' performance by chance. Specifically, the out-of-sample predictions produced by the BBC-CVT (249 predictions) were permuted (sampled without replacement) 5,000 times and the models' performance statistics (coefficient of determination) were estimated for each permutation. This null distribution was then used to assess the observed model performance statistics in the non-permuted data. That is, the p-value for testing the significance of models' performance was determined by computing the proportion of resampled statistics at least as high as 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 models' whole-sample explanatory power, the Akaike Information Criterion (AIC) for each model was obtained. The AIC balances the goodness-of-fit of the model (model accuracy) against its complexity (the number of estimable parameters of the model) so that a 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 (1-2 units), strong (2-4 units), considerably strong (4-7 units) and decisive (>10) evidence against evidence against M_1 (Burnham & Anderson, 2016). We used the AIC difference to compare the explanatory power of alternative models of the same response variable (cognitive or demographic) with the same connectivity modality as predictor (SC, FC or CC). The models under comparison only differed in the type of parcellation used to compute the connectivity predictors. Consequently, if two parcellations had similar explanatory power but one was coarser than the other, the coarser parcellation was favoured due to requiring fewer parameters (number of connectivity predictors) to model the same response data.

To compare the models' out-of-sample predictive performance, the models' out-of-sample predictions (generated during repeated BBC-CVT) were used to obtain a bootstrapped (5000 samples) estimates of the models' performance statistics (coefficient of determination). Additionally, the non-parametric Wilcoxon rank sum test for equal medians was used to assess the significance of differences in performance between different models. These comparisons were only done for models which performed better than chance.

Graph theory measures

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. Small world propensity requires computation of the clustering coefficient and the shortest path length of the network (Muldoon et al., 2016). The clustering coefficient was obtained using Onnela's algorithm (Muldoon et al., 2016; Onnela et al., 2005) whereas the shortest path length was obtained using the Floyd–Warshall Algorithm applied to the weighted graph obtained from the inverse of each connectivity matrix. Modularity statistic requires computation of network modules, which were defined with Newman's algorithm (Newman, 2006).

To study the impact of the parcellation on global network organisation, a series of paired t-tests were used to test for significant differences between the same global graph theoretic measure computed for two different parcellation schemes. To explore the impact of global network architecture yielded by each parcellation on predictive modelling, regression models were fitted to each observed cognitive domain using graph theory. Graph theory models were fitted and compared in the same manner as for connectivity but without elastic-net regularisation.

RESULTS

Connectivity

Predictive modelling of demographics

AIC-based model comparison

AIC values demonstrated that across all connectivity modalities and parcellations, age was best modelled with SC defined either with AAL3 (166) or Shen (93) parcellations (Figure 1). The difference between AIC values of SC models defined with either AAL3 (166) or Shen (93) was small. For SC, there was evidence in favour of both AAL3 (166) and Shen (93) parcellations relative to Shen (184) and Shen (278), and Brainnetome (246). For FC, there was substantial evidence in favour of Shen (93) parcellation to AAL3 (166). For CC, there was substantial evidence in favour of Shen (93) parcellation above to Shen (184).

Next, across connectivity modalities and parcellations, education had lowest AIC value when modelled with SC defined with Shen (93). Within SC, there was evidence in favour of Shen (93) relative to AAL3 (166), Shen (278) and Brainnetome (246) and substantial evidence in favour of Shen (184). For FC, Brainnetome (246) was favoured to Shen (93). For CC, Shen (93) was favoured above Shen (184) and Shen (278).

Finally, there was no evidence to suggest that sex could be modelled more effectively using any single connectivity modality or parcellation.

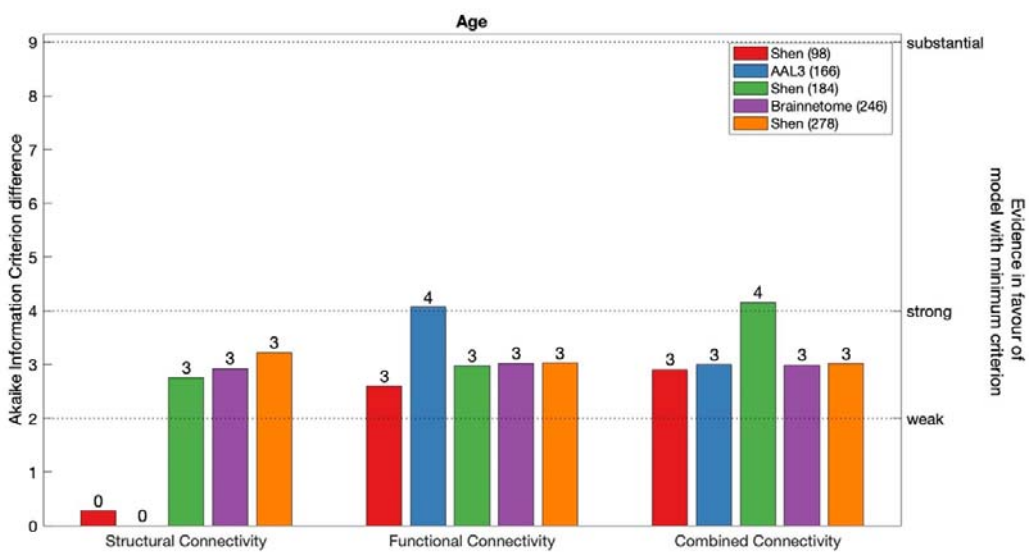

Figure 1 AIC difference for SC, FC and CC models of age, constructed with each parcellation scheme. Dotted line marks substantial evidence in favour of CC model defined with Shen (246) (minimum AIC across all modalities and parcellations).

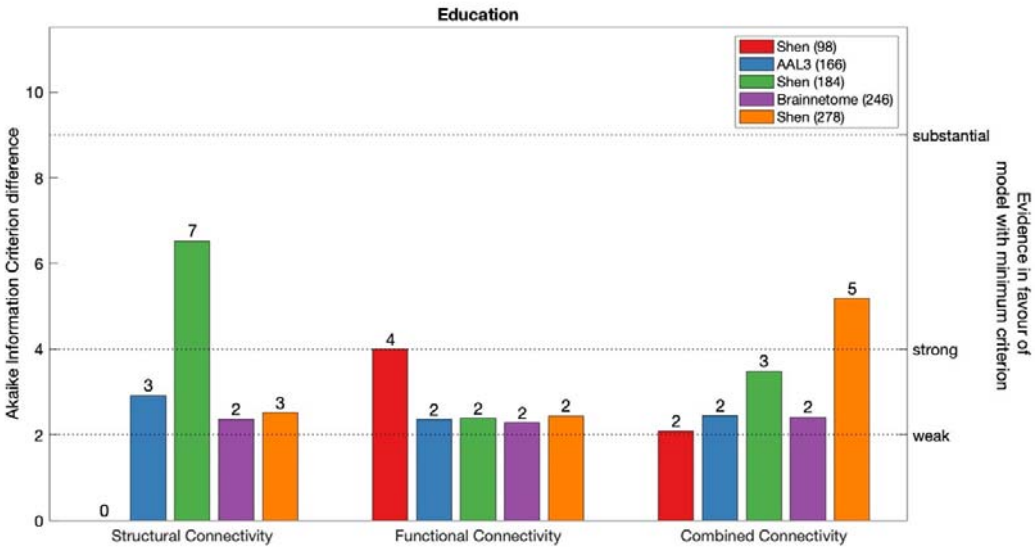

Figure 2 AIC difference for SC, FC and CC models of education, constructed with each parcellation scheme. Dotted line marks substantial evidence in favour of CC model defined with Shen (246) (minimum AIC across all modalities and parcellations).

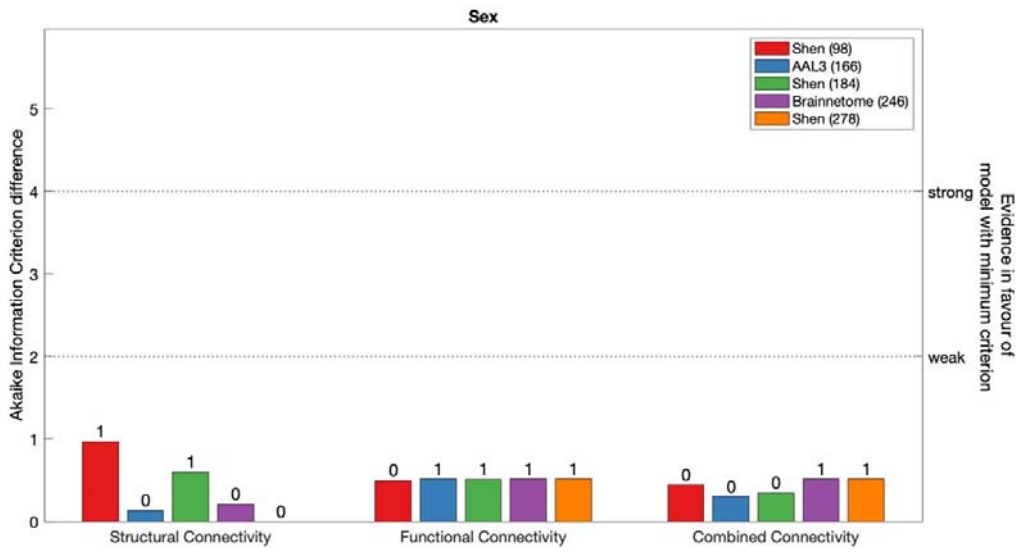

Figure 3 AIC difference for SC, FC and CC models of sex, constructed with each parcellation scheme. Dotted line marks substantial evidence in favour of CC model defined with Shen (246) (minimum AIC across all modalities and parcellations).

Model comparison based on out-of-sample performance

Results of the cross-validation of models of demographics are presented in Figures 4-6. Cross-validation results demonstrated that only FC defined using Shen (278) parcellation did not produce above chance out-of-sample predictions of Education. All remaining models of all demographic variables performed above chance regardless of the parcellation schemes or the connectivity modality used as predictor.

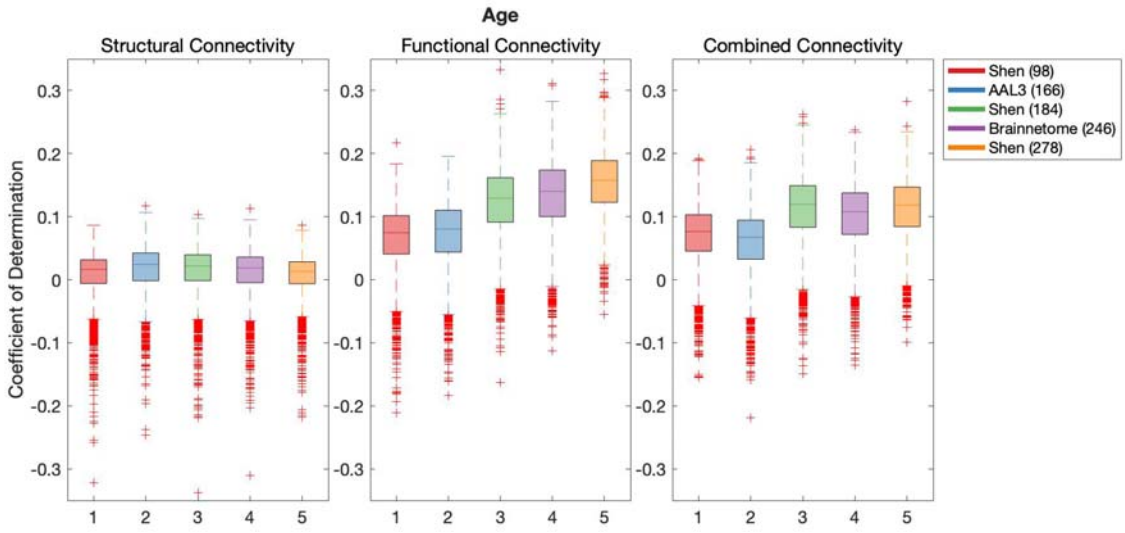

Figure 4 Results of BBC-CV of age models constructed with SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate above- chance predictive performance and unfilled boxes illustrate bellow-chance prediction.

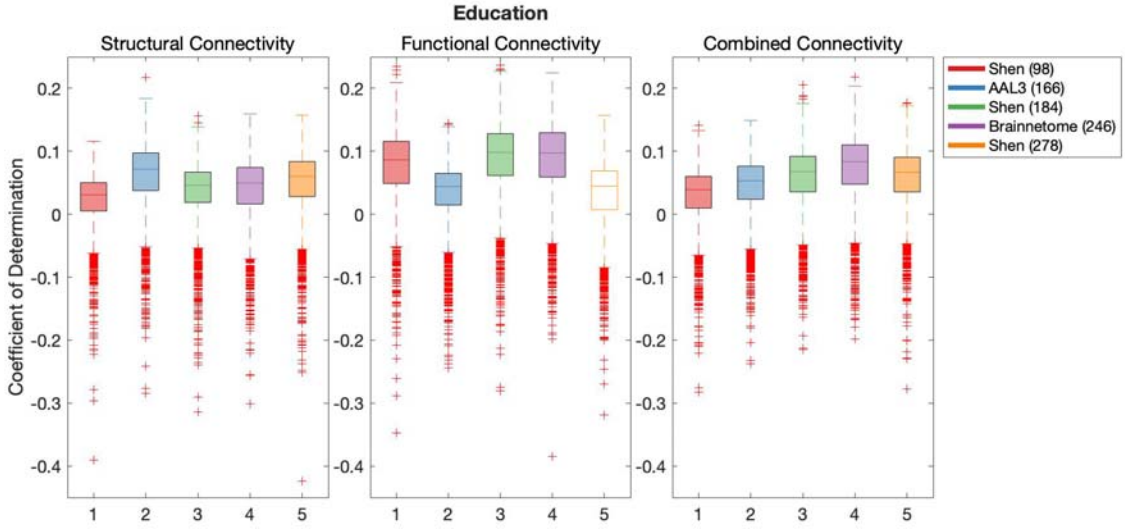

Figure 5 Results of BBC-CV of education models constructed with SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate above- chance predictive performance and unfilled boxes illustrate bellow-chance prediction.

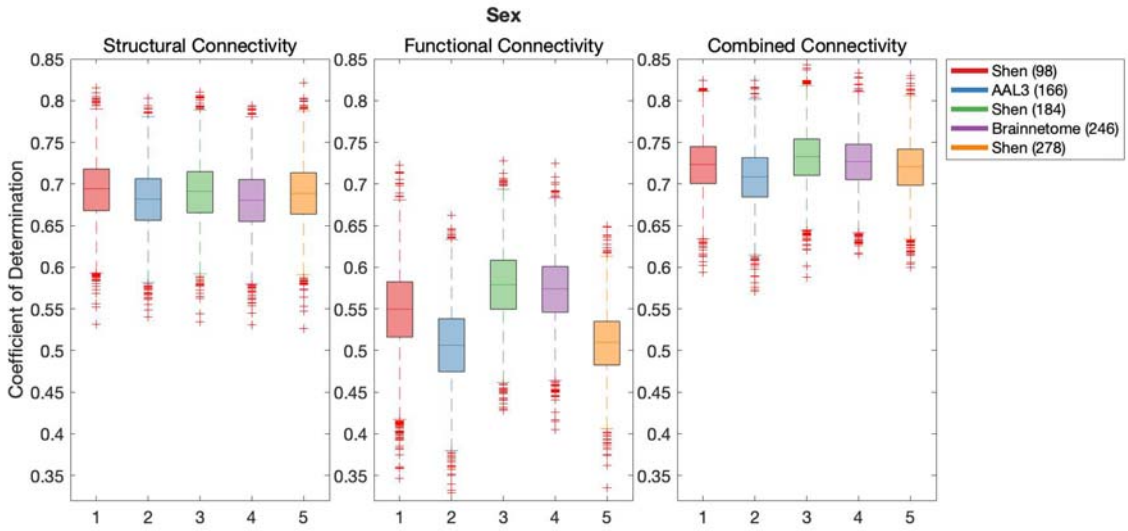

Figure 6 Results of BBC-CV of sex models constructed with SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate above- chance predictive performance and unfilled boxes illustrate bellow-chance prediction.

Results of Wilcoxon rank sum test comparing out-of-sample models' performance are shown in Table 1, where the False Discovery Rate (FDR) adjusted significance threshold equals 0.0095 (Benjamini & Hochberg, 1995)). The following results and effect sizes (Rosenthal et al., 1994) have been written for significant pair-wise comparisons presented in Table 1.

Out-of-sample predictive performance of age was significantly higher for models using AAL3 (166) parcellation in SC than any other parcellation, for Shen (278) in FC and for Shen (184) in CC. In the case of education, out-of-sample predictive performance was also significantly higher for models using AAL3 (166) parcellation in SC than alternative parcellations. In FC, highest out-of-sample predictive performance for education was found for both Shen (184) and Brainnetome (246), and there was no significant difference between their prediction performance. For CC, Brainnetome (246) parcellation performed more effectively at predicting education than alternative parcellations. Finally, Sex was most effectively predicted using Shen (93) parcellation in SC, and Shen (184) in both FC and CC.

Table 1 Pairwise comparison Z-scores of coefficients of determination of demographics generated during BBC-CVT. 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 meet the FDR-corrected p-value of 0.0095.

Age

Connectivity		Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Structural	Shen (93)	-	**-14.07**	**-10.6**	**-4.97**	**4.66**
	AAL3 (166)	**14.07**	-	**3.76**	**9.27**	**18.47**
	Shen (184)	**10.6**	**-3.76**	-	**5.65**	**15.22**
	Brainnetome (246)	**4.97**	**-9.27**	**-5.65**	-	**9.61**
	Shen (278)	**-4.66**	**-18.47**	**-15.22**	**-9.61**	-
Functional	Shen (93)	-	**-6.76**	**-50.22**	**-55.7**	**-70.37**
	AAL3 (166)	**6.76**	-	**-44.4**	**-50.66**	**-65.95**
	Shen (184)	**50.22**	**44.4**	-	**-10.03**	**-28.36**
	Brainnetome (246)	**55.7**	**50.66**	**10.03**	-	**-17.78**
	Shen (278)	**70.37**	**65.95**	**28.36**	**17.78**	-
Combined	Shen (93)	-	**10.93**	**-41.68**	**-32.3**	**-42.57**
	AAL3 (166)	**-10.93**	-	**-48.91**	**-40.62**	**-50.07**
	Shen (184)	**41.68**	**48.91**	-	**11.47**	**0.62**
	Brainnetome (246)	**32.3**	**40.62**	**-11.47**	-	**-11.14**
	Shen (278)	**42.57**	**50.07**	**-0.62**	**11.14**	-

Education

Connectivity		Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Structural	Shen (93)	-	**-45.49**	**-21.13**	**-24.15**	**-35.74**
	AAL3 (166)	**45.49**	-	**30.57**	**24.61**	**13.74**
	Shen (184)	**21.13**	**-30.57**	-	**-5.43**	**-18.15**
	Brainnetome (246)	**24.15**	**-24.61**	**5.43**	-	**-11.95**
	Shen (278)	**35.74**	**-13.74**	**18.15**	**11.95**	-
Functional	Shen (93)	-	**43.94**	**-11.91**	**-11.59**	
	AAL3 (166)	**-43.94**	-	**-52.78**	**-51.19**	-
	Shen (184)	**11.91**	**52.78**	-	0.02	-
	Brainnetome (246)	**11.59**	**51.19**	-0.02	-	-
	Shen (278)	-	-	-	-	-
Combined	Shen (93)	-	**-19.49**	**-34.74**	**-46.51**	**-33.69**
	AAL3 (166)	**19.49**	-	**-17.51**	**-32.87**	**-16.18**
	Shen (184)	**34.74**	**17.51**	-	**-17.27**	1.49
	Brainnetome (246)	**46.51**	**32.87**	**17.27**	-	**18.7**
	Shen (278)	**33.69**	**16.18**	-1.49	**-18.7**	-

Connectivity		Sex				
		Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Structural	Shen (93)	-	**15.09**	**3.45**	**16.85**	**5.98**
	AAL3 (166)	**-15.09**	-	**-11.97**	1.91	**-9.36**
	Shen (184)	**-3.45**	**11.97**	-	**13.8**	**2.6**
	Brainnetome (246)	**-16.85**	-1.91	**-13.8**	-	**-11.2**
	Shen (278)	**-5.98**	**9.36**	**-2.6**	**11.2**	-
Functional	Shen (93)	-	**40.03**	**-30.34**	**-25.6**	**41.02**
	AAL3 (166)	**-40.03**	-	**-63.69**	**-61.65**	**-3.16**
	Shen (184)	**30.34**	**63.69**	-	**6.52**	**66.43**
	Brainnetome (246)	**25.6**	**61.65**	**-6.52**	-	**64.62**
	Shen (278)	**-41.02**	**3.16**	**-66.43**	**-64.62**	-
Combined	Shen (93)	-	**21.11**	**-14.08**	**-4.77**	**4.12**
	AAL3 (166)	**-21.11**	-	**-33.98**	**-25.74**	**-17.36**
	Shen (184)	**14.08**	**33.98**	-	**9.53**	**18.19**
	Brainnetome (246)	**4.77**	**25.74**	**-9.53**	-	**8.96**
	Shen (278)	**-4.12**	**17.36**	**-18.19**	**-8.96**	-

Predictive modelling of cognition

AIC-based model comparison

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

AIC favoured models of Executive Function composed with FC defined with AAL3 (166) parcellation. For SC, there was substantial evidence in favour of Brainnetome (246) parcellation. For both FC and CC, there was substantial evidence in favour of AAL3 (166) parcellation.

There was substantial evidence in favour of FC model of Self-regulation constructed with AAL3 (166). In addition, for SC, there was substantial evidence in favour of Shen (184) parcellation. For CC, there was substantial evidence in favour of AAL3 (166) parcellation.

Overall, Language was best modelled by CC defined with Shen (184) parcellation. Similarly, when FC was used to model Language it was also best defined with Shen (184) parcellation. However, when FS was used to model Language it was best defined with AAL3 (166) parcellation.

Overall, Encoding was best modelled with FC defined with Shen (278) parcellation. When SC was used to model Encoding, best models were achieved with AAL3 (166) and Shen (93). When CC was used to model Encoding, according to AIC differences, the preferred model was defined with Shen (184).

Finally, across connectivity modalities and parcellations, Sequence Processing was best modelled with Brainnetome (246). When either SC or CC was used, AIC values were in favour of models obtained with AAL3 (166).

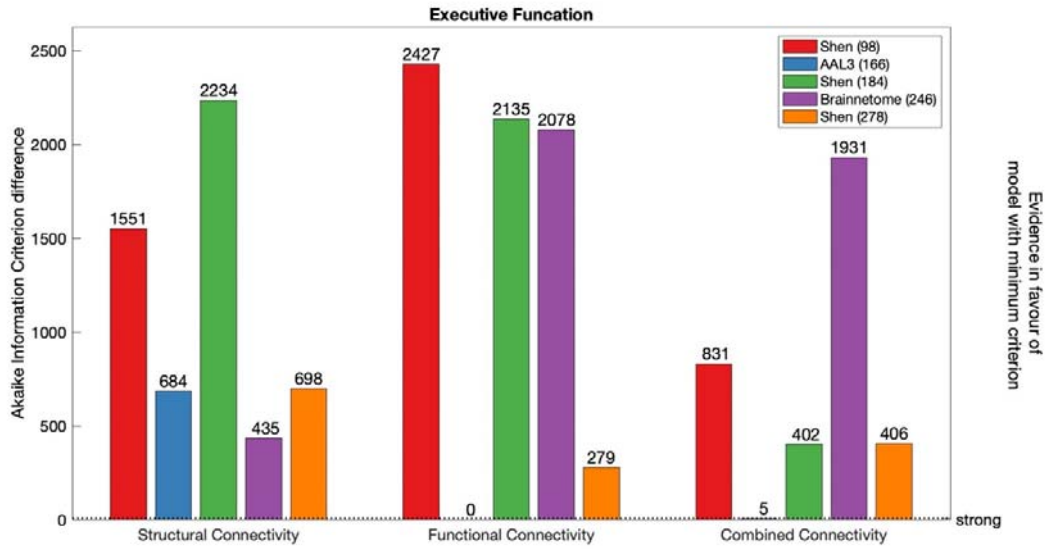

Figure 7 AIC difference for SC, FC and CC models of Executive Function, constructed with each parcellation scheme. Dotted line marks strong evidence in favour of FC model defined with AAL3 (166) (minimum AIC across all modalities and parcellations).

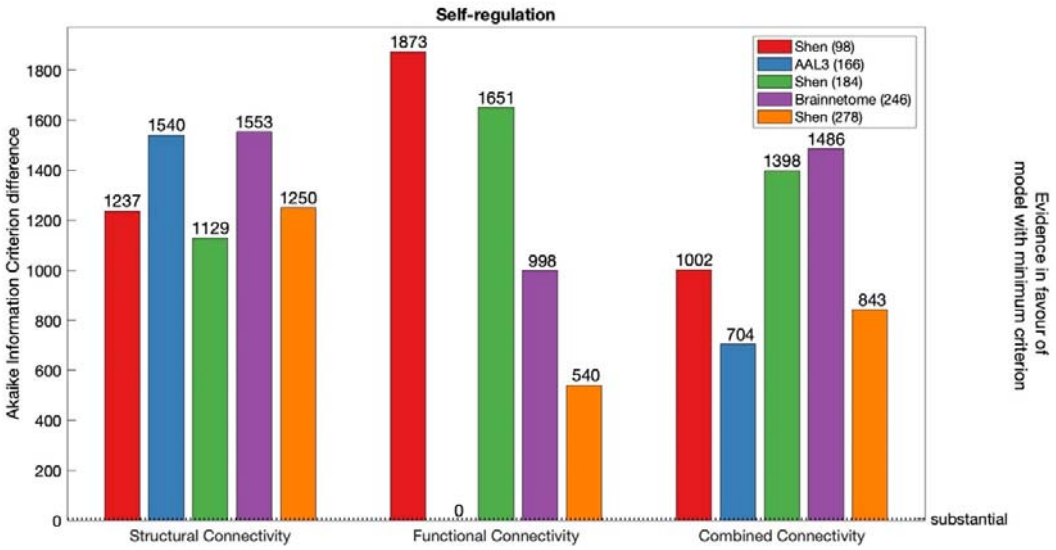

Figure 8 AIC difference for SC, FC and CC models of Self-regulation, constructed with each parcellation scheme. Dotted line marks substantial evidence in favour of FC model defined with AAL3 (166) (minimum AIC across all modalities and parcellations).

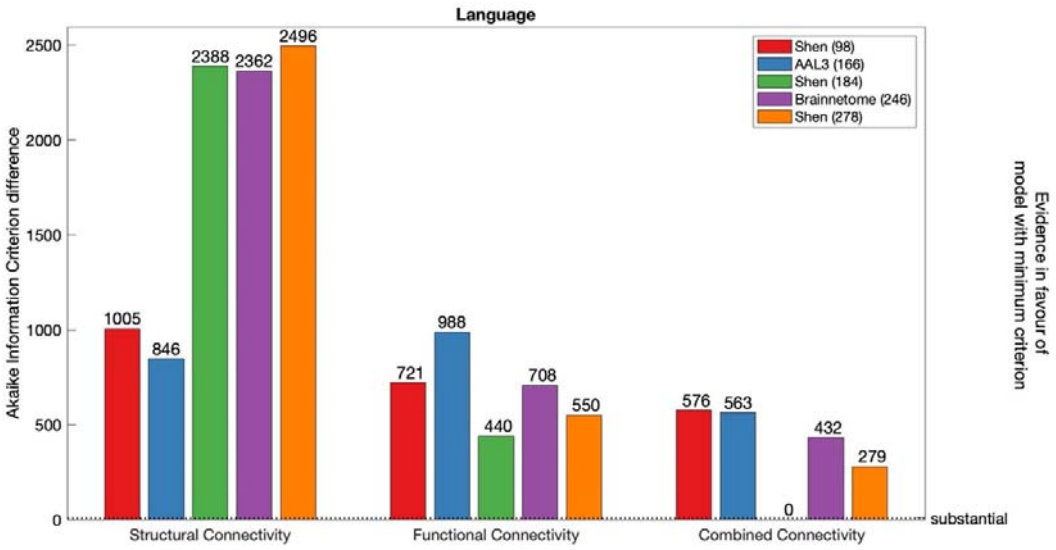

Figure 9 AIC difference for SC, FC and CC models of Language, constructed with each parcellation scheme. Dotted line marks substantial evidence in favour of CC model defined with Shen (246) (minimum AIC across all modalities and parcellations).

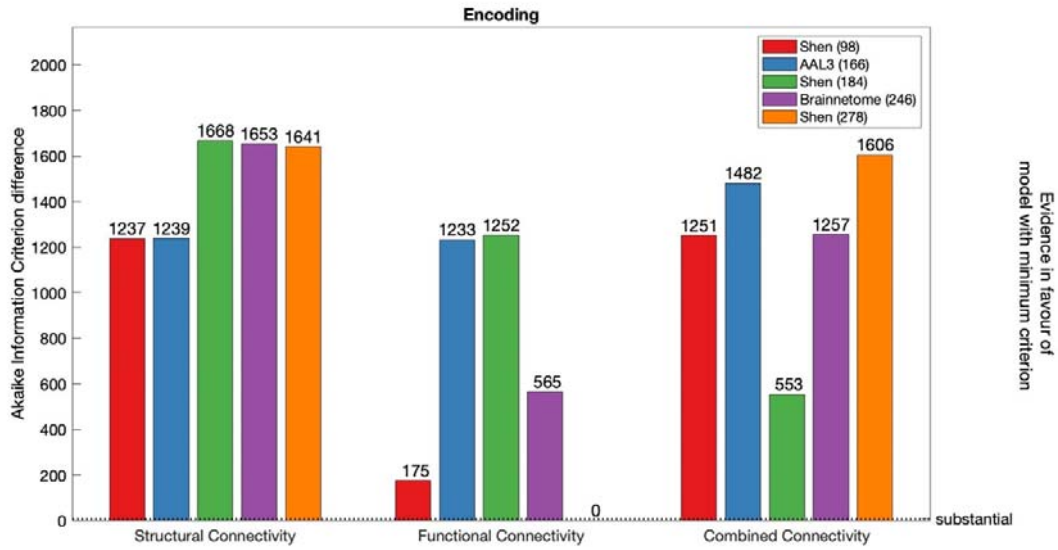

Figure 10 AIC difference for SC, FC and CC models of Encoding, constructed with each parcellation scheme. Dotted line marks substantial evidence in favour of FC model defined with Shen (278) (minimum AIC across all modalities and parcellations).

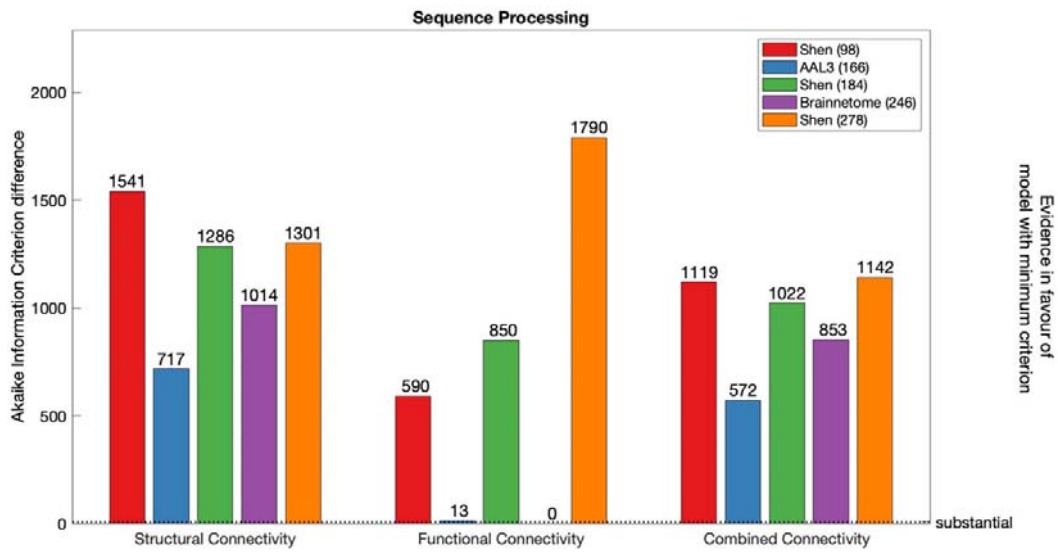

Figure 11 AIC difference for SC, FC and CC models of Encoding, constructed with each parcellation scheme. Dotted line marks substantial evidence in favour of FC model defined with Brainnetome (246) (minimum AIC across all modalities and parcellations).

Model comparison based on out-of-sample performance

Results of the cross-validation of models of cognition are presented in Figures 12-17. It was found that the Shen (93) parcellation consistently produced out-of-sample predictions of cognition that explained more variance in unseen samples than chance. This was observed across SC, FC and CC.

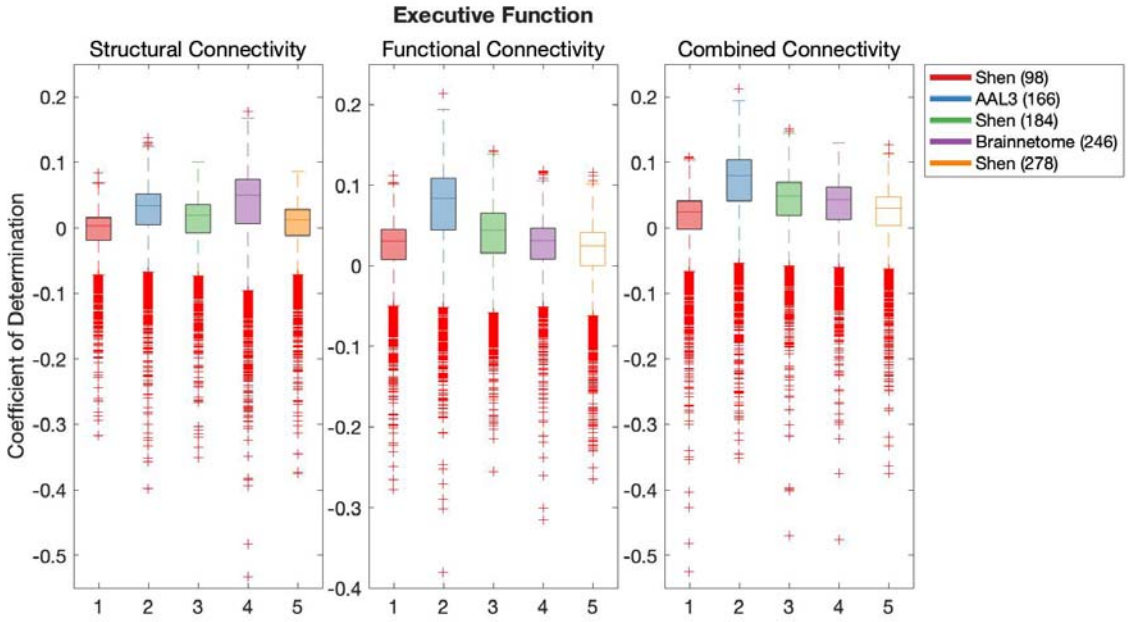

Figure 12 Results of BBC-CV of Executive Function models constructed with SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate above- chance predictive performance and unfilled boxes illustrate bellow-chance prediction.

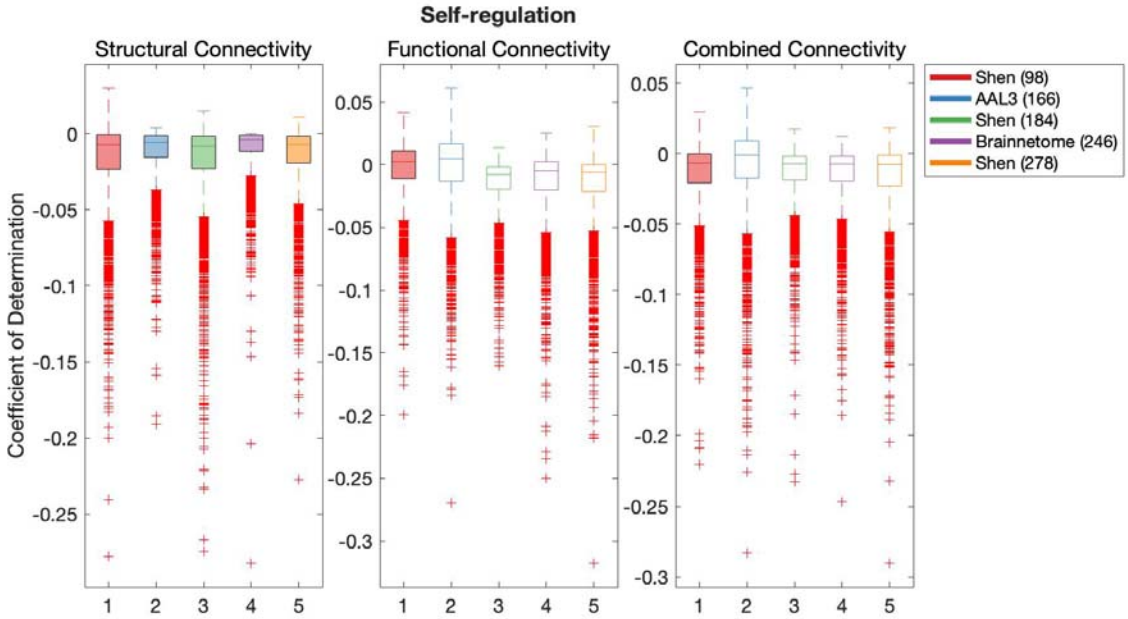

Figure 13 Results of BBC-CV of Self-regulation models constructed with SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate above- chance predictive performance and unfilled boxes illustrate bellow-chance prediction.

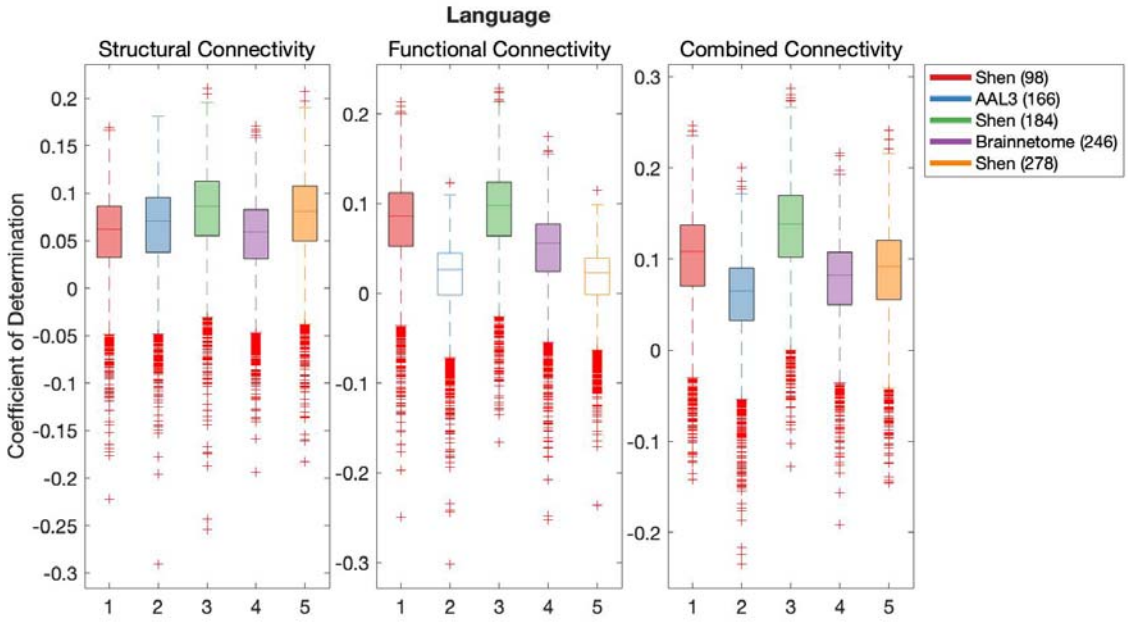

Figure 14 Results of BBC-CV of Language models constructed with SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate above- chance predictive performance and unfilled boxes illustrate bellow-chance prediction.

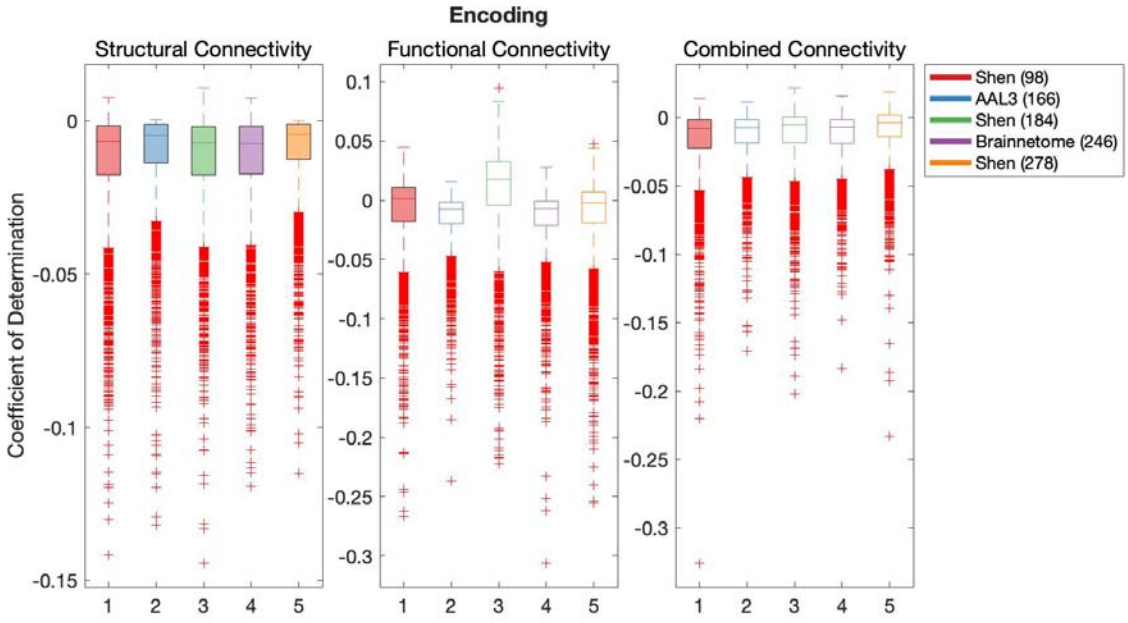

Figure 15 Results of BBC-CV of Encoding models constructed with SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate above- chance predictive performance and unfilled boxes illustrate bellow-chance prediction.

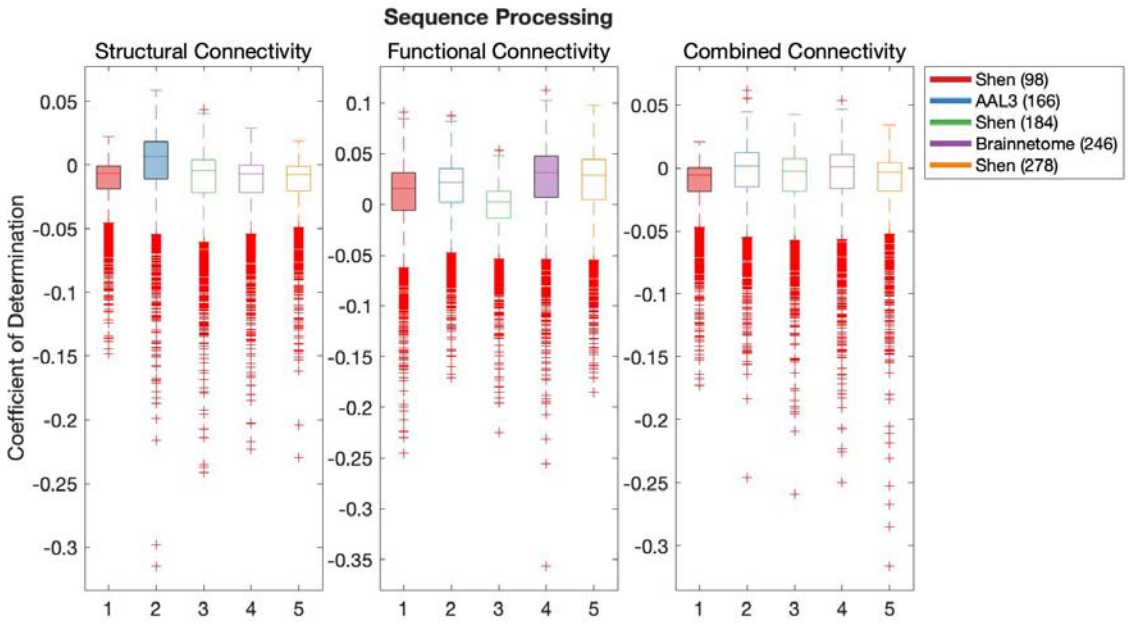

Figure 16 Results of BBC-CV of Sequence Processing models constructed with SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Filled boxes illustrate above- chance predictive performance and unfilled boxes illustrate bellow-chance prediction.

All pairwise comparisons of the cross-validation performance 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.

Despite consistently explaining more variance in unseen sample than chance, Shen (93) rarely outperformed other parcellations at producing effective predictions of cognition (vs. Shen 184 Self-regulation structural connectivity, Language vs Brainnetome (246) structural connectivity, functional connectivity and combined connectivity, Language vs AAL3 (166) with CC and Language vs Shen 278 with CC). Next, AAL3 (166) consistently produced the most effective predictions of Executive Function utilising FC and CC. Similarly, Brainnetome (246) consistently produced the most effective predictions of Self-regulation utilising SC. Shen (184) parcellation consistently produced the most effective predictions of Language utilising FC. Shen (278) parcellation consistently produced the most effective predictions of Encoding utilising SC. There was no other parcellation scheme that consistently outperformed other parcellation schemes for any given type of connectivity or any cognitive construct.

Table 2 Pairwise comparison Z-scores of coefficients of determination of cognition generated during BBC-CVT. 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 <0.0184

		Executive Function				
Connectivity		Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Structural	Shen (93)	-	**-43.85**	**-27.88**	**-46.45**	**-18.63**
	AAL3 (166)	**43.85**	-	**21.77**	**-19.14**	**31.34**
	Shen (184)	**27.88**	**-21.77**	-	**-33.76**	**11.09**
	Brainnetome (246)	**46.45**	**19.14**	**33.76**	-	**39.49**
	Shen (278)	**18.63**	**-31.34**	**-11.09**	**-39.49**	-
Functional	Shen (93)	-	**-52.55**	**-22.66**	-2.36	-
	AAL3 (166)	**52.55**	-	**39.15**	**51.1**	-
	Shen (184)	**22.66**	**-39.15**	-	**20.2**	-
	Brainnetome (246)	2.36	**-51.1**	**-20.2**	-	-
Combined	Shen (93)	-	**-52.73**	**-33.88**	**-26.44**	-
	AAL3 (166)	**52.73**	-	**32.49**	**38.96**	-
	Shen (184)	**33.88**	**-32.49**	-	**8.85**	-
	Brainnetome (246)	**26.44**	**-38.96**	**-8.85**	-	-
		Self-regulation				
		Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Structural	Shen (93)	-	-2.28	**5.46**	**-9.21**	**2.86**
	AAL3 (166)	2.28	-	**8.78**	**-8.56**	**6.21**
	Shen (184)	**-5.46**	**-8.78**	-	**-16.65**	**-2.95**
	Brainnetome (246)	**9.21**	**8.56**	**16.65**	-	**14.65**
	Shen (278)	**-2.86**	**-6.21**	**2.95**	**-14.65**	-
		Language				
		Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Structural	Shen (93)	-	**9.33**	**-19.47**	**-13.5**	**12.71**
	AAL3 (166)	**-9.33**	-	**-28.35**	**-22.5**	**3.39**
	Shen (184)	**19.47**	**28.35**	-	**5.93**	**31.57**

	Brainnetome (246)	**13.5**	**22.5**	**-5.93**	-	**25.76**
	Shen (278)	**-12.71**	**-3.39**	**-31.57**	**-25.76**	-
Functional	Shen (93)	**-**	-	**-12.74**	**35.59**	-
	Shen (184)	**12.74**	**-**	-	**45.9**	-
	Brainnetome (246)	**-35.59**	**-**	**-45.9**	-	-
Combined	Shen (93)	-	**42.56**	**-30.72**	**-20.46**	**15.58**
	AAL3 (166)	**-42.56**	-	**-63.98**		**-29.41**
	Shen (184)	**30.72**	**63.98**	-	**52.63**	**43.99**
	Brainnetome (246)	**-26.16**	**20.46**	**-52.63**	-	**-10.74**
	Shen (278)	**-15.58**	**29.41**	**-43.99**	**10.74**	-
				Encoding		
		Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Structural	Shen (93)	-	**-7.24**	1.98	0.97	**-9.45**
	AAL3 (166)	**7.24**	-	**9.79**	**8.57**	-1.98
	Shen (184)	-1.98	**-9.79**	-	-0.87	**-12.1**
	Brainnetome (246)	-0.97	**-8.57**	0.87	-	**-10.91**
	Shen (278)	**9.45**	1.98	**12.1**	**10.91**	-
				Sequence Processing		
		Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Structural	Shen (93)	-	**38.51**	-	-	-
	AAL3 (166)	**-38.51**	-	-	-	-
Functional	Shen (93)	-	-	-	**-25.86**	-
	Brainnetome (246)	**25.86**	-	-	-	-

Graph theory

Comparison of global organisation across parcellations

Figures 17 and 18 show the distribution of scores obtained for each participant's global graph theory measures for each parcellation. 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, structural parcellation scheme (AAL3, 166) produced high modularity statistic in SC, but low modularity statistic in FC. Further, high resolution scheme (Shen, 278) produced lowest

global efficiency and transitivity in SC, whereas in FC a lower resolution parcellation (AAL3, 166) produced lowest global efficiency and transitivity.

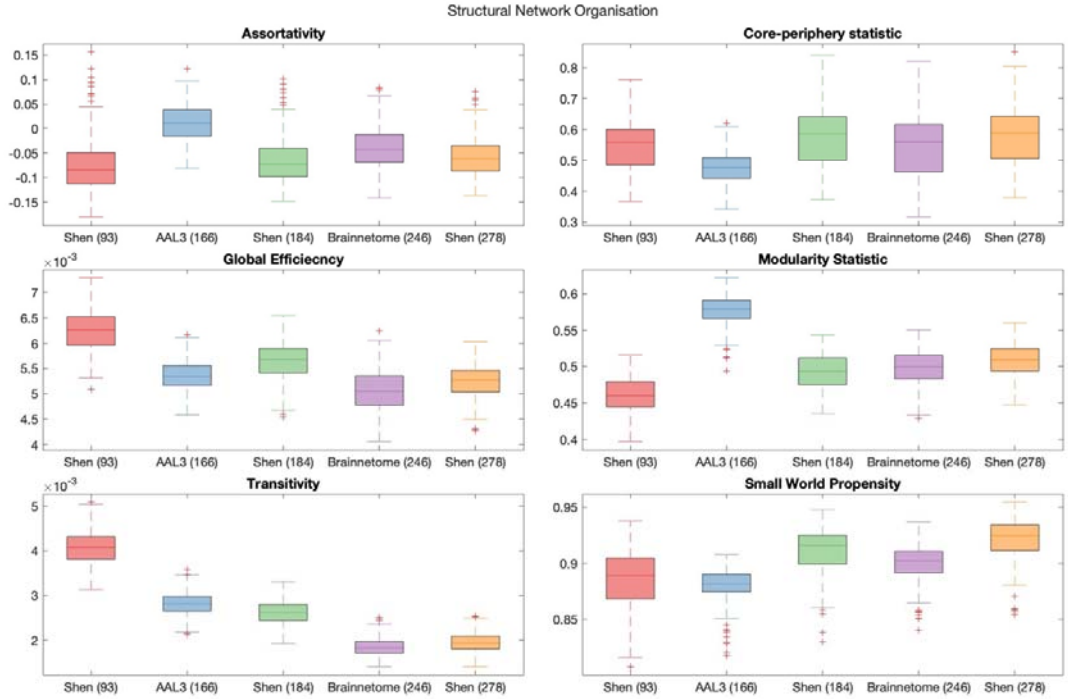

Figure 17 Graph theory measures of the global network organisation of SC.

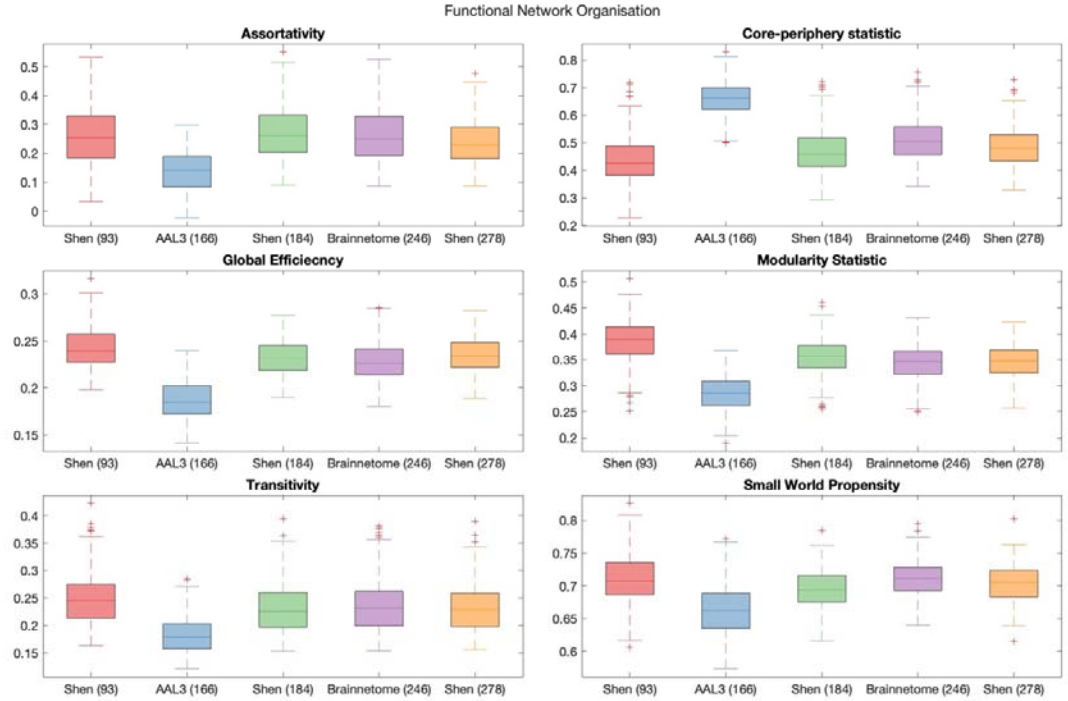

Figure 18 Graph theory measures of the global network organisation of FC.

Table 3 Pairwise comparison t-statistics of structural connectivity organisation 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 has higher global property (assortativity, core-periphery statistic, global efficiency, modularity, transitivity, small world propensity) than the parcellation presented in a given column, whereas negative values indicate that the parcellation presented in a given row has lower global property than the parcellation presented in a given column. All t-scores presented in bold have passed the FDR-adjusted critical p-value =<.0275.

			Assortativity		
	Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Shen (93)	-	**-25.16**	**-6.53**	**-14.84**	**-8.27**
AAL3 (166)	**25.16**	-	**24.66**	**19.27**	**24.26**
Shen (184)	**6.53**	**-24.66**	-	**-13.65**	**-6.94**
Brainnetome (246)	**14.84**	**-19.27**	**13.65**	-	**10.22**
Shen (278)	**8.27**	**-24.26**	**6.94**	**-10.22**	-
			Coreness Statistic		
	Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Shen (93)	-	**23.34**	**-34.67**	-0.75	**-34.91**
AAL3 (166)	**-23.34**	-	**-28.32**	**-19.58**	**-29.75**
Shen (184)	**34.67**	**28.32**	-	**17.61**	**-6.91**
Brainnetome (246)	0.75	**19.58**	**-17.61**	-	**-21.08**
Shen (278)	**34.91**	**29.75**	**6.91**	**21.08**	-
			Global Efficiency		
	Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Shen (93)		**51.98**	**82.15**	**71.26**	**108.86**
AAL3 (166)	**-51.98**		**-17.69**	**20.23**	**5.48**
Shen (184)	**-82.15**	**17.69**		**32.85**	**93.32**
Brainnetome (246)	**-71.26**	**-20.23**	**-32.85**		**-10.62**
Shen (278)	**-108.86**	**-5.48**	**-93.32**	**10.62**	
			Modularity statistic		
	Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Shen (93)		**-63.76**	**-32.99**	**-29.8**	**-47.5**
AAL3 (166)	**63.76**		**47.31**	**43.92**	**39.43**
Shen (184)	**32.99**	**-47.31**		**-5.07**	**-19.69**
Brainnetome (246)	**29.8**	**-43.92**	**5.07**		**-9.32**
Shen (278)	**47.5**	**-39.43**	**19.69**	**9.32**	
			Transitivity		
	Shen (93)	AAL3	Shen	Brainnetome	Shen

		(166)	(184)	(246)	(278)
Shen (93)		**68.75**	**202.09**	**121.69**	**208.09**
AAL3 (166)	**-68.75**		**11.51**	**75.16**	**61.78**
Shen (184)	**-202.09**	**-11.51**		**56.06**	**183.45**
Brainnetome (246)	**-121.69**	**-75.16**	**-56.06**		**-7.82**
Shen (278)	**-208.09**	**-61.78**	**-183.45**	**7.82**	

Small World Propensity

	Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Shen (93)		**3.98**	**-31.98**	**-11.32**	**-37.98**
AAL3 (166)	**-3.98**		**-29.09**	**-19.21**	**-43.74**
Shen (184)	**31.98**	**29.09**		**12.1**	**-22.83**
Brainnetome (246)	**11.32**	**19.21**	**-12.1**		**-26.27**
Shen (278)	**37.98**	**43.74**	**22.83**	**26.27**	

Table 3 Pairwise comparison t-statistics of structural connectivity organisation 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 has higher global property (assortativity, core-periphery statistic, global efficiency, modularity, transitivity, small world propensity) than the parcellation presented in a given column, whereas negative values indicate that the parcellation presented in a given row has lower global property than the parcellation presented in a given column. All t-scores presented in bold have passed the FDR-adjusted critical p-value =<.0275.

Assortativity

	Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Shen (93)	-	**20.88**	**-2.63**	-0.09	**6.39**
AAL3 (166)	**-20.88**	-	**-21.62**	**-21.62**	**-17.77**
Shen (184)	**2.63**	**21.62**	-	**3**	**13.9**
Brainnetome (246)	0.09	**21.62**	**-3**	-	**8.8**
Shen (278)	**-6.39**	**17.77**	**-13.9**	**-8.8**	-

Core-Periphery Statistic

	Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Shen (93)	-	**-68.37**	**-18.33**	**-32.32**	**-21.85**
AAL3 (166)	**68.37**	-	**60.28**	**52.11**	**59.06**
Shen (184)	**18.33**	**-60.28**	-	**-27.03**	**-12.64**
Brainnetome (246)	**32.32**	**-52.11**	**27.03**	-	**17.08**
Shen (278)	**21.85**	**-59.06**	**12.64**	**-17.08**	-

Global Efficiency

	Shen	AAL3	Shen	Brainnetome	Shen

	(93)	(166)	(184)	(246)	(278)
Shen (93)	-	**84.7**	**22.21**	**31.2**	**14.9**
AAL3 (166)	**-84.7**	-	**-75.97**	**-70.77**	**-76.49**
Shen (184)	**-22.21**	**75.97**	-	**10.37**	**-15.42**
Brainnetome (246)	**-31.2**	**70.77**	**-10.37**	-	**-17.25**
Shen (278)	**-14.9**	**76.49**	**15.42**	**17.25**	-
			Modularity statistic		
	Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Shen (93)	-	**56.89**	**21.18**	**27.29**	**23.92**
AAL3 (166)	**-56.89**	-	**-42.09**	**-37.11**	**-37.5**
Shen (184)	**-21.18**	**42.09**	-	**9.3**	**7.55**
Brainnetome (246)	**-27.29**	**37.11**	**-9.3**	-	**-2.52**
Shen (278)	**-23.92**	**37.5**	**-7.55**	**2.52**	-
			Transitivity		
	Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Shen (93)	-	**51.56**	**22.37**	**15.01**	**19.87**
AAL3 (166)	**-51.56**	-	**-41.06**	**-40.5**	**-42.83**
Shen (184)	**-22.37**	**41.06**	-	**-7.33**	-0.46
Brainnetome (246)	**-15.01**	**40.5**	**7.33**	-	**6.14**
Shen (278)	**-19.87**	**42.83**	0.46	**-6.14**	-
			Small World Propensity		
	Shen (93)	AAL3 (166)	Shen (184)	Brainnetome (246)	Shen (278)
Shen (93)	-	**18**	**8.4**	0.69	**3.94**
AAL3 (166)	**-18**	-	**-15.9**	**-24.86**	**-20.78**
Shen (184)	**-8.4**	**15.9**	-	**-12.39**	**-8.88**
Brainnetome (246)	-0.69	**24.86**	**12.39**	-	**5.67**
Shen (278)	**-3.94**	**20.78**	**8.88**	**-5.67**	-

Predictive modelling of demographics

AIC-based model comparison

Figures 19-21 demonstrate AIC for models of demographics composed with global graph theory measures of SC, FC and CC when defined with the 5 parcellation schemes.

AIC values demonstrated that across all connectivity modalities and parcellations, age was best modelled with FC defined Brainnetome (246) parcellation (Figure 19). For SC, there was weak evidence in favour of Shen (93 and 184) and Brainnetome (246) relative to AAL3 (166). For FC, there was weak evidence in favour of Brainnetome (246) parcellation to AAL3 (166) and Shen (278) and strong-to-substantial evidence in favour of Brainnetome (246) to Shen (93 and 184). For CC, there was strong-to-substantial evidence in favour of Brainnetome (246) parcellation to all alternatives.

Next, across connectivity modalities and parcellations, education had lowest AIC value when modelled with SC defined with Brainnetome (246) or with FC defined with Shen (93 and 184) parcellation. Within SC, there was strong evidence in favour of Brainnetome (246) relative to alternative parcellations. For FC, there was no notable difference between AIC values of models defined with Shen (184) and Shen (93). For CC, there was weak-to-substantial evidence in favour of Brainnetome (246) relative to alternative parcellations.

Finally, there was substantial evidence in favour of SC and CC models of sex relative FC models. For SC and CC, lowest AIC values were estimated for SC models defined with either Shen (184) followed by Shen (278). These two models were substantially favoured to any other alternative models.

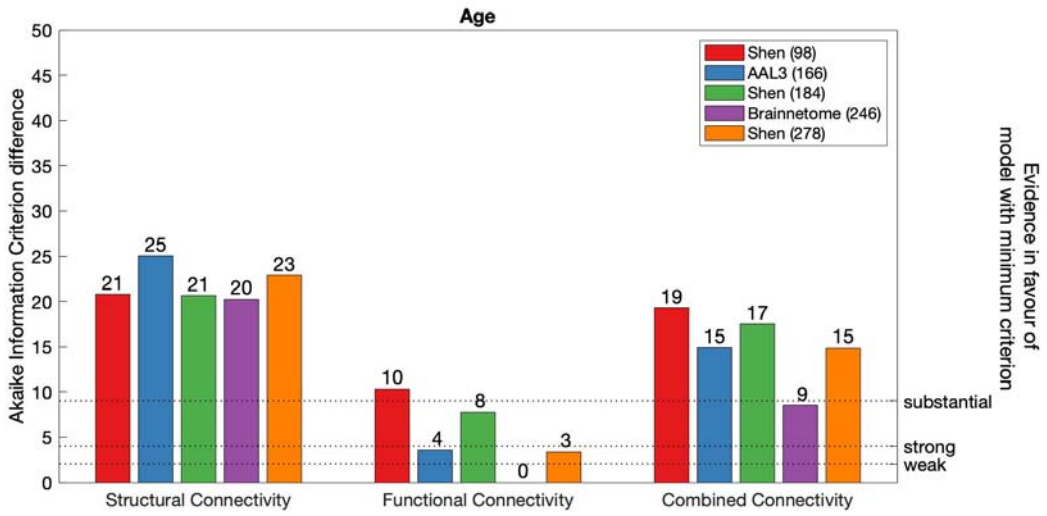

Figure 19 AIC difference for global graph theory measure models of Age, constructed with each parcellation scheme. Dotted line marks weak, strong and substantial evidence in favour of FC model defined with Brainnetome (246) (minimum AIC across all modalities and parcellations).

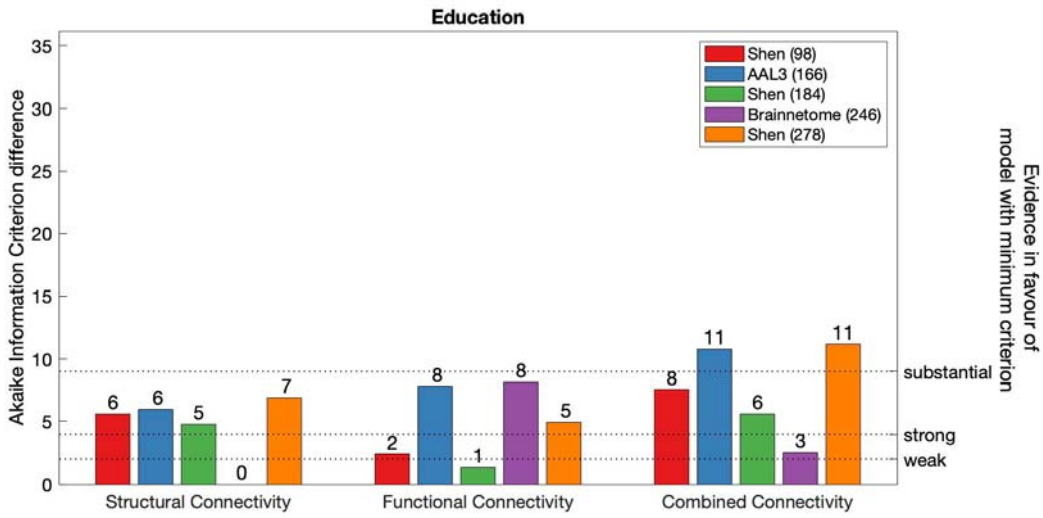

Figure 20 AIC difference for global graph theory measure models of Education, constructed with each parcellation scheme. Dotted line marks weak, strong and substantial evidence in favour of SC model defined with Brainnetome (246) (minimum AIC across all modalities and parcellations).

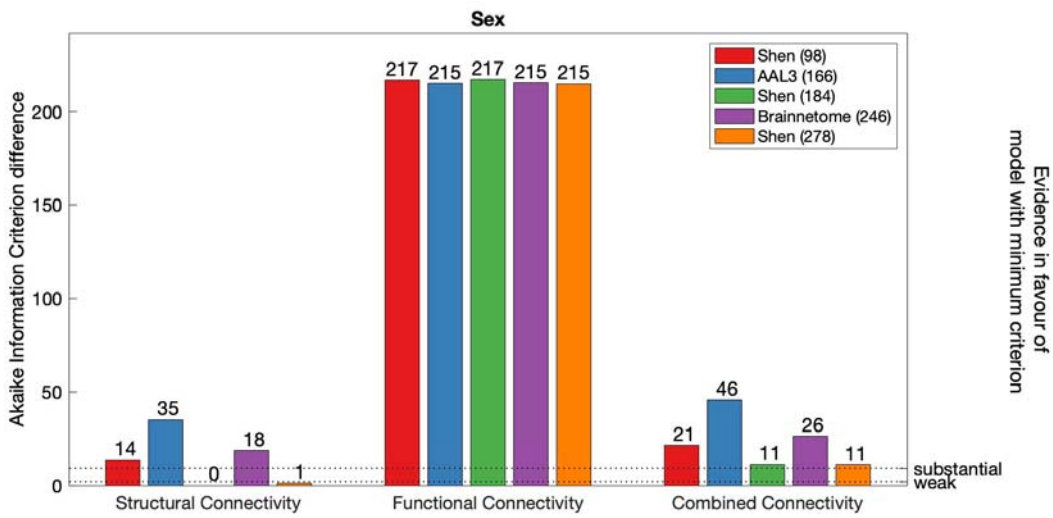

Figure 21 AIC difference for global graph theory measure models of Sex, constructed with each parcellation scheme. Dotted line marks weak, and substantial evidence in favour of SC model defined with Shen (184) (minimum AIC across all modalities and parcellations).

Model comparison based on out-of-sample performance

Results of the cross-validation of models of demographics are presented in Figures 22-24. It was found that graph theory measures of global organisation of SC, FC and CC could not produce out-of-sample predictions of demographics that would explain more variance in unseen samples than chance.

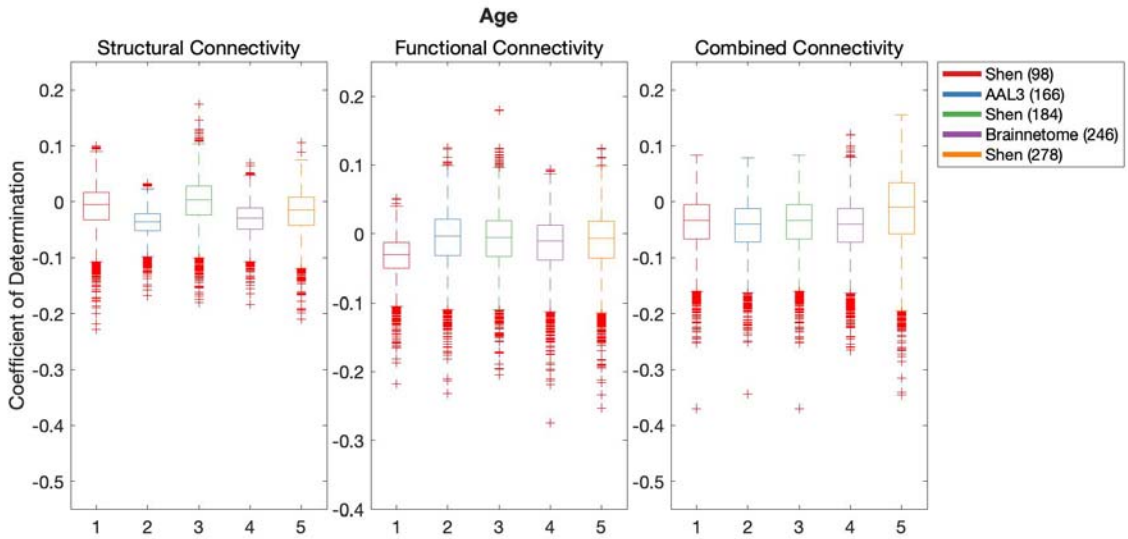

Figure 22 Results of BBC-CV of age models constructed with graph theory measures of SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Unfilled boxes illustrate bellow-chance prediction

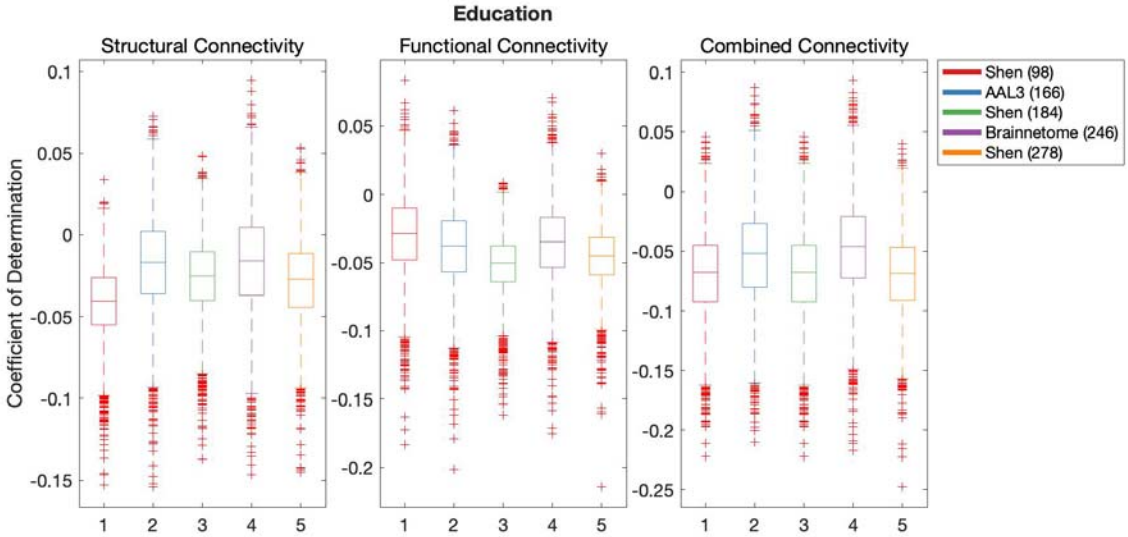

Figure 23 Results of BBC-CV of education models constructed with graph theory measures of SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Unfilled boxes illustrate bellow-chance prediction

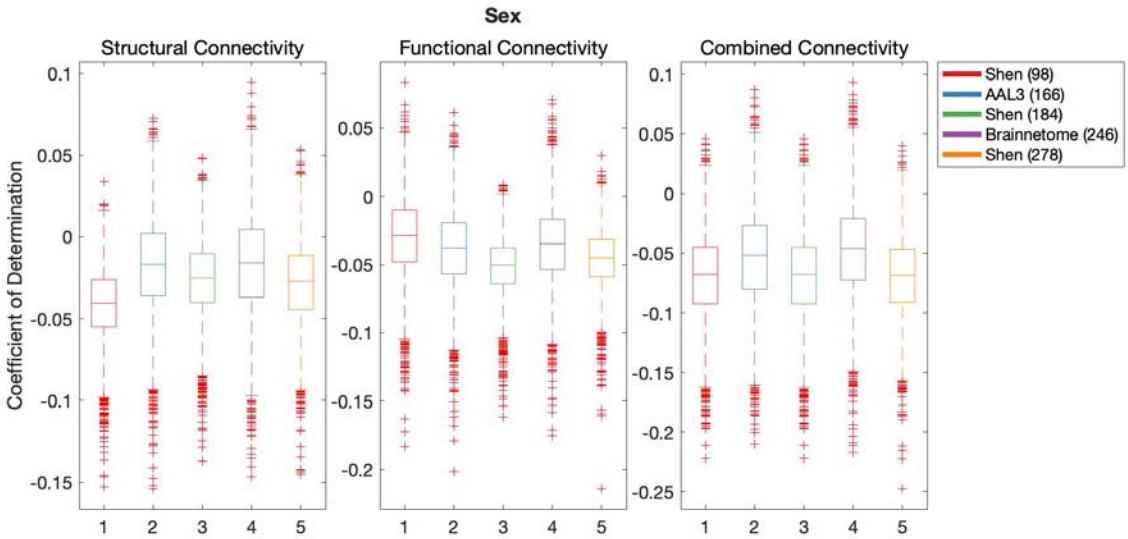

Figure 24 Results of BBC-CV of sex models constructed with graph theory measures of SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Unfilled boxes illustrate bellow-chance prediction

Predictive modelling of cognition

AIC-based model comparison

Figures 25-29 demonstrate AIC for models of cognition composed with global graph theory measures of SC, FC and CC when defined with the 5 parcellation schemes.

Overall, AIC favoured models of Executive Function composed with SC defined with Shen (184) parcellation. In addition, there was no notable difference between SC, FC and CC models of Executive Function defined with Shen (184) parcellation. For SC, FC and CC, there was strong-to-substantial evidence in favour of Shen (184) relative to alternative parcellations.

Self-regulation was best modelled by SC when parcellated with Shen (278). For SC, there was no notable difference between AIC values of Shen (278) and Shen (93). However, there was weak evidence in favour of Shen (278) relative to Shen (184) and Brainnetome (246), and strong evidence relative to AAL3 (166). For FC, Self-regulation was best modelled with AAL3 (166) but the evidence in favour of this model relative to Shen (93 and 278) was not notable, and it was weak relative to Shen (184) and Brainnetome (246). In For CC, there was no notable difference between AIC values of Shen (93) and Shen (278). However, there was strong evidence in favour of Shen (278) relative to AAL3 (166) and Shen (184), and substantial evidence relative to Brainnetome (246).

Language was best modelled by SC defined with Shen (184) parcellation. For SC, there was no notable difference between AIC values of Shen (184) and Shen (278), but there was strong to substantial evidence in favour of Shen (184) relative to alternative parcellations. For FC, AIC difference weakly favoured AAL3 (166) parcellation relative to Shen (93, 184), and strongly relative to Brainnetome (246). For CC, there was no notable

difference between AIC values of Shen (278) and Shen (184) and AAL3 (166) but there was strong-to-substantial evidence in favour of these models relative to Shen (93) and Brainnetome (246).

Overall, Encoding was best modelled with SC defined with Shen (93) parcellation. When FC was used to model Encoding, best models were achieved with either AAL3 (166) or Brainnetome (246). When CC was used to model Encoding, according to AIC differences, the preferred model was defined with AAL3 (93) and Shen (93).

Finally, Sequence Processing was best modelled with FC defined with either Shen (93) or Shen (184). When either SC or CC was used, AIC values were in favour of models obtained with Shen (184).

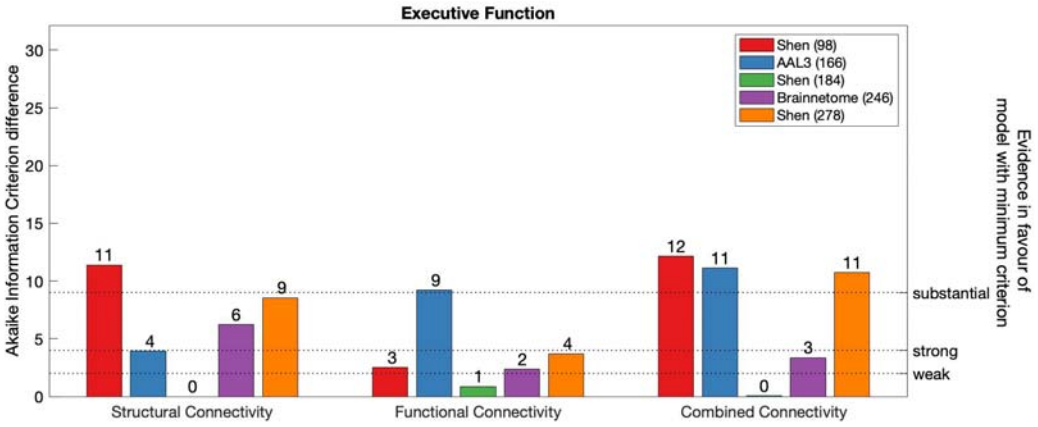

Figure 25 AIC difference for global graph theory measure models of Executive Function, constructed with each parcellation scheme. Dotted line marks weak, and substantial evidence in favour of SC model defined with Shen (184) (minimum AIC across all modalities and parcellations).

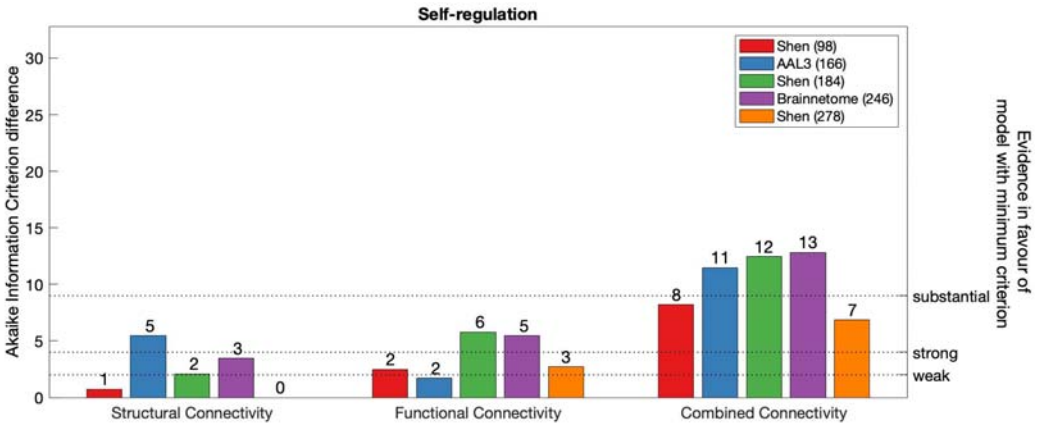

Figure 26 AIC difference for global graph theory measure models of Self-regulation, constructed with each parcellation scheme. Dotted line marks weak, and substantial evidence in favour of SC model defined with Shen (278) (minimum AIC across all modalities and parcellations).

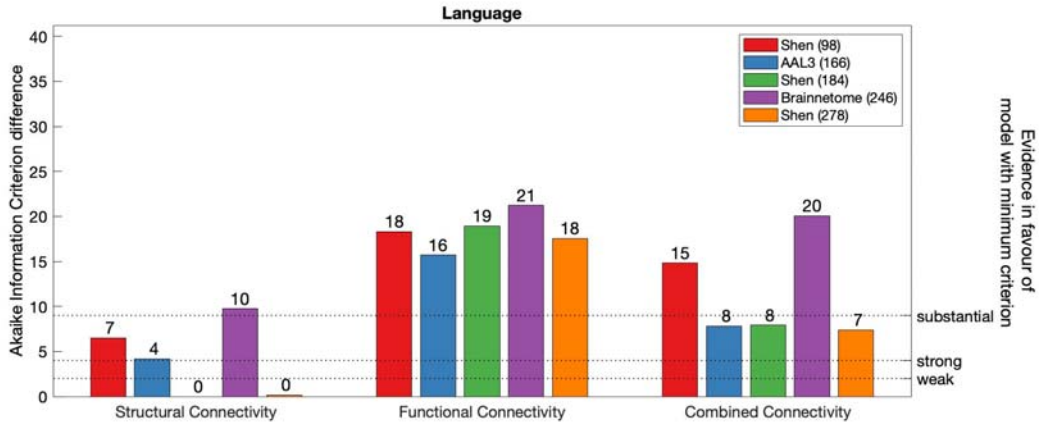

Figure 27 AIC difference for global graph theory measure models of Language, constructed with each parcellation scheme. Dotted line marks weak, and substantial evidence in favour of SC model defined with Shen (184) (minimum AIC across all modalities and parcellations).

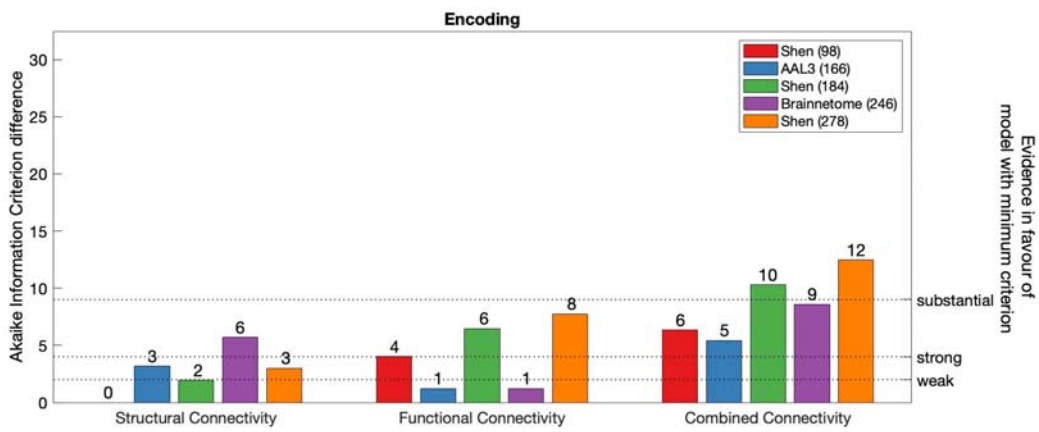

Figure 28 AIC difference for global graph theory measure models of Encoding, constructed with each parcellation scheme. Dotted line marks weak, and substantial evidence in favour of SC model defined with Shen (93) (minimum AIC across all modalities and parcellations).

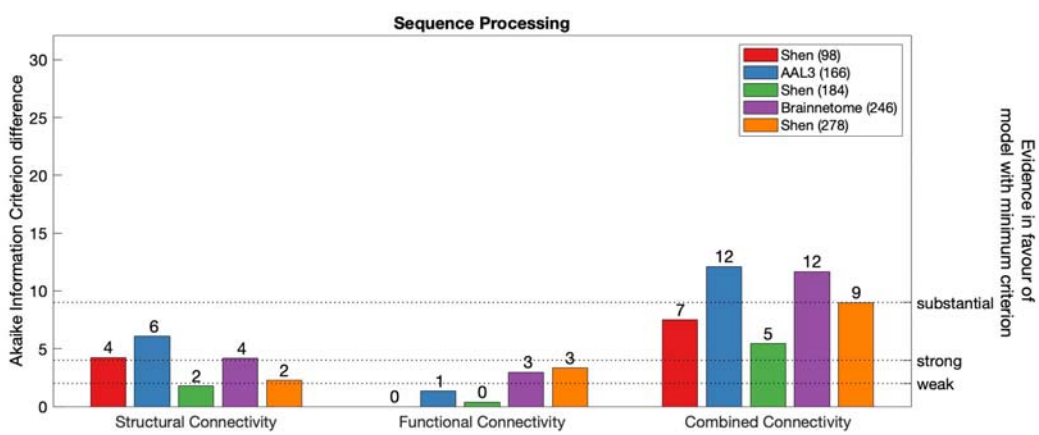

Figure 29 AIC difference for global graph theory measure models of Sequence Processing, constructed with each parcellation scheme. Dotted line marks weak, and substantial evidence in favour of SC model defined with Shen (93) (minimum AIC across all modalities and parcellations).

Model comparison based on out-of-sample performance

Results of the cross-validation of models of cognition are presented in Figures 30-34. It was found that graph theory measures of global organisation of SC, FC and CC could not produce out-of-sample predictions of cognition that would explain more variance in unseen samples than chance.

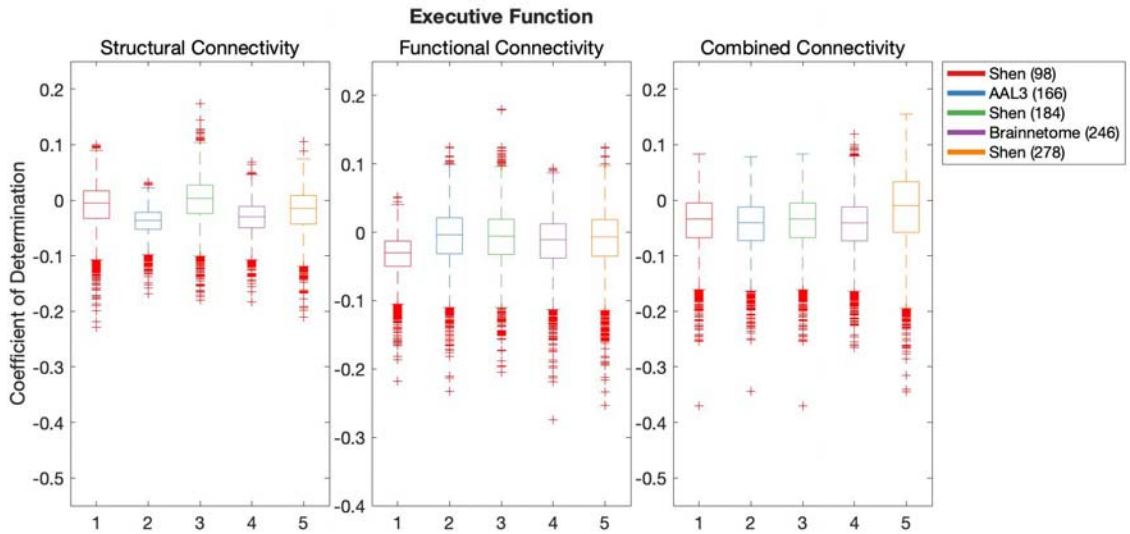

Figure 30 Results of BBC-CV of Executive Function models constructed with graph theory measures of SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Unfilled boxes illustrate bellow-chance prediction

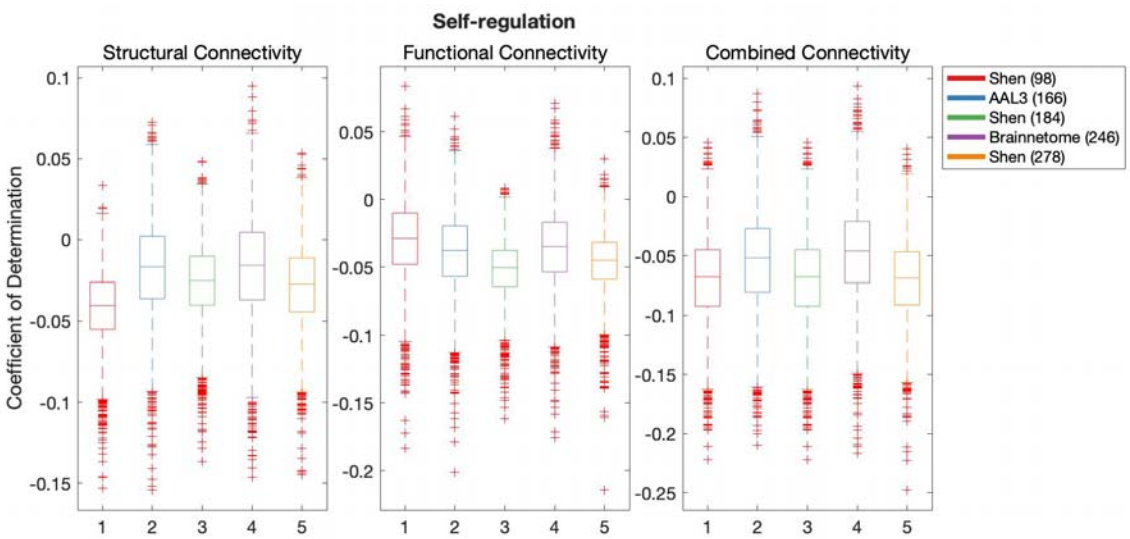

Figure 31 Results of BBC-CV of Self-regulation models constructed with graph theory measures of SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Unfilled boxes illustrate bellow-chance prediction

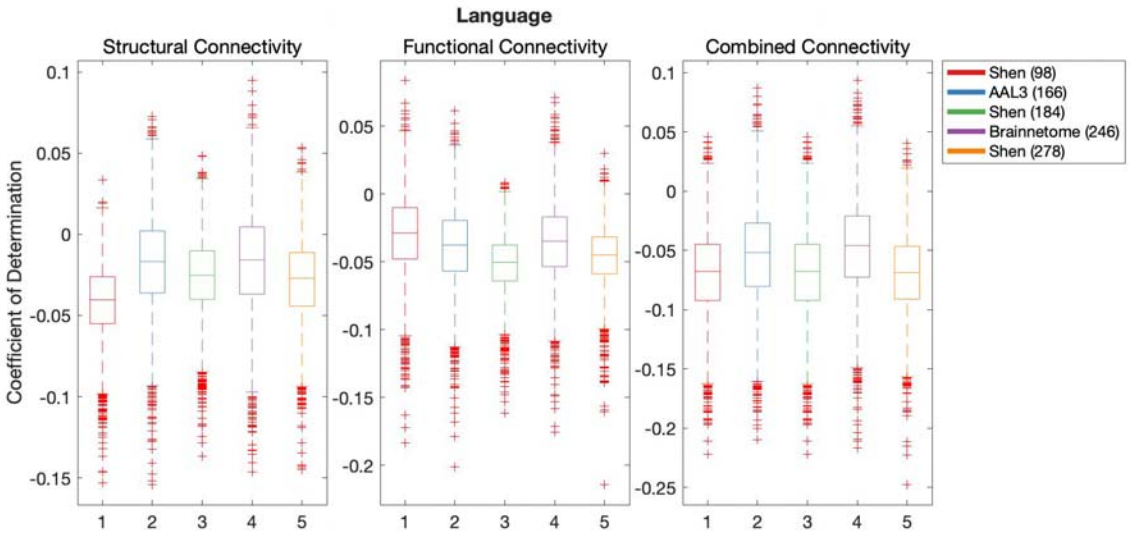

Figure 32 Results of BBC-CV of Language models constructed with graph theory measures of SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Unfilled boxes illustrate bellow-chance prediction

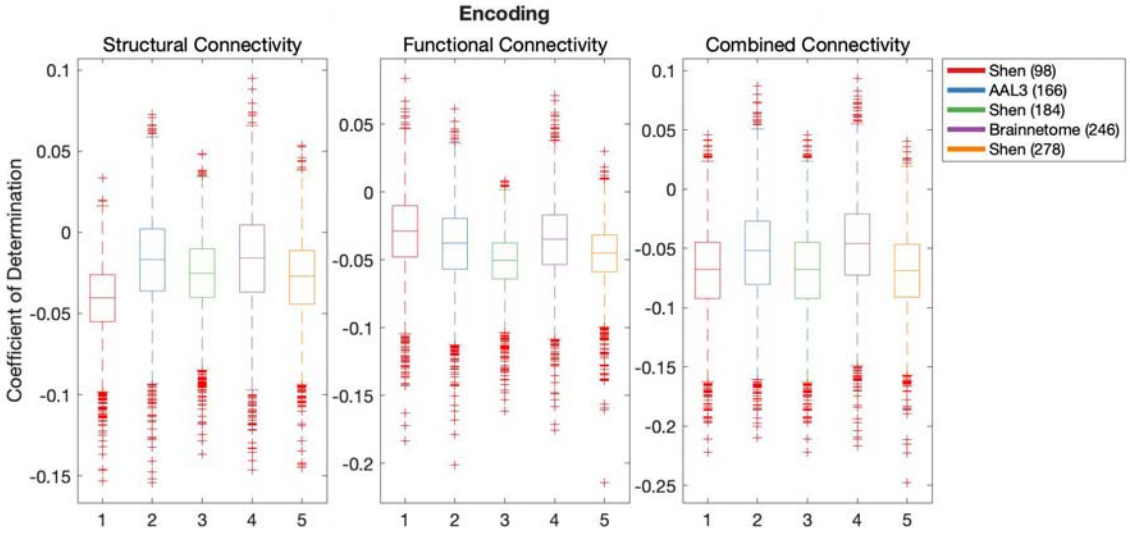

Figure 33 Results of BBC-CV of Encoding models constructed with graph theory measures of SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Unfilled boxes illustrate bellow-chance prediction

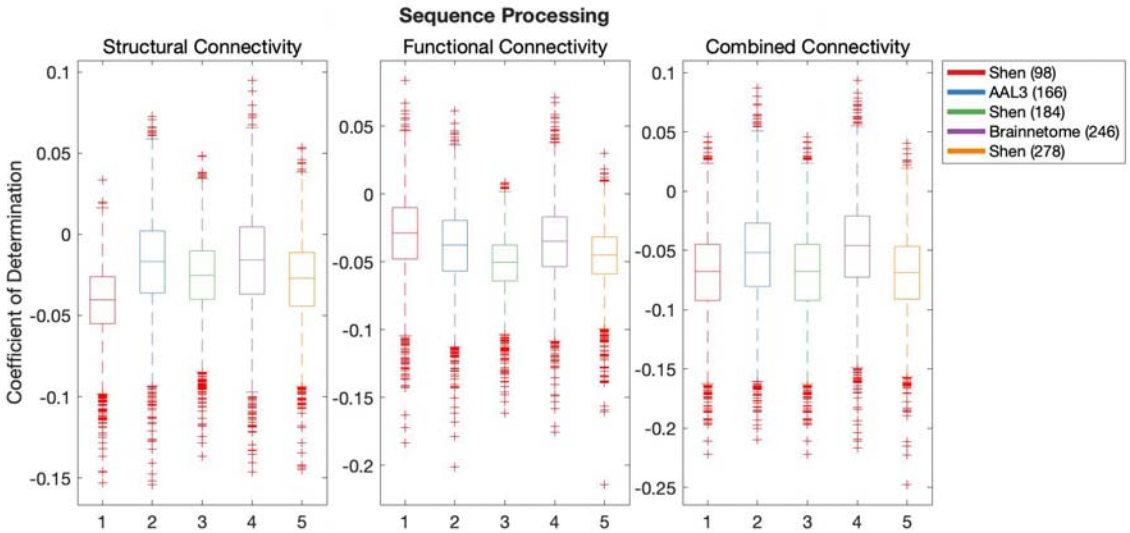

Figure 34 Results of BBC-CV of Sequence Processing models constructed with graph theory measures of SC, FC and CC, as measured by the coefficient of determination. The solid lines show the median scores, the boxes show the interquartile range (IQR), and ticks outside of whiskers indicate outlier scores across all bootstrap samples. Unfilled boxes illustrate bellow-chance prediction

DISCUSSION

Currently available brain parcellation schemes vary in the neuroimaging data and the node-generation algorithms used to obtain them. Consequently, the obtained parcellations vary in the number, size and shape of nodes. These differences 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 is systematically impacting predictive models. To fill this gap in literature, the present research defined SC, FC and CC with 5 different popular parcellation schemes and constructed predictive models of demographics and cognition based on models using each connectivity modality and parcellation scheme as predictors. Results of AIC-based model comparison did not show 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 93) to model cognition. Analysis of the out-of-sample performance of the models showed that most parcellations produced above-chance predictive performance of demographics for all connectivity modalities. The only exception to this was the model of educational level using FC with the Shen 278 parcellation. In addition, it was found that models of demographics showed higher performance when using FC defined with functional parcellations. There was no consistent benefit to using low or high-resolution parcellations for this connectivity modality. When SC was used to model age and education, AAL3 (166) parcellation produced the highest out-of-sample performance. In contrast, models of sex based on SC modality benefitted from using low-resolution functional parcellation (Shen, 93). When out-of-sample predictions of cognition were assessed, it was found that no parcellation scheme would consistently produce significantly higher predictive performance than the others,

even when the connectivity and parcellation corresponded to the same modality (functional or structural). However, out-of-sample predictions of cognition were consistently above chance when using connectivity defined with the Shen 93 parcellation scheme. 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 (as reflected by out-of-sample predictions), when FC was defined with AAL3 (166) 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), which suggests that smooth parcel boundaries may better describe connectivity of the system. This need for overlapping parcels 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, 93) 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.

A key finding of this work was that although parcellation schemes impacted the global architecture of brain networks, the global architecture of the networks had no out-of-sample predictive power with demographics or cognition. 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 demographics or 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 of the present work is that 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 whether 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 93) 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 93 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.*

Burnham, K. P., & Anderson, D. R. (2016). Multimodel Inference. *Sociological Methods & Research*, *33*(2), 261-304. https://doi.org/10.1177/0049124104268644

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

Hilger, K., Ekman, M., Fiebach, C. J., & Basten, U. (2017). Efficient hubs in the intelligent brain: Nodal efficiency of hub regions in the salience network is associated with general intelligence. *Intelligence*, *60*, 10-25. https://doi.org/10.1016/j.intell.2016.11.001

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

Langer, N., Pedroni, A., Gianotti, L. R. R., Hänggi, J., Knoch, D., & Jäncke, L. (2012). Functional brain network efficiency predicts intelligence. *Human Brain Mapping*, *33*(6), 1393-1406. https://doi.org/10.1002/hbm.21297

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

Rosenthal, R., Cooper, H., & Hedges, L. (1994). Parametric measures of effect size. *The handbook of research synthesis*, *621*(2), 231-244.

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