Shared functional connections within and between cortical networks predict individual cognitive abilities in males and females

Running Title: Shared features predict cognition across sexes

Elvisha Dhamala[1,2,*], Keith W. Jamison[1], Abhishek Jaywant[3-5], Amy Kuceyeski[1,2,*]

[1] Department of Radiology, Weill Cornell Medicine, New York, NY, USA, 10065

[2] Brain and Mind Research Institute, Weill Cornell Medicine, New York, NY, USA, 10065

[3] Department of Psychiatry, Weill Cornell Medicine, New York, NY, USA, 10065

[4] Department of Rehabilitation Medicine, Weill Cornell Medicine, New York, NY, USA, 10065

[5] NewYork-Presbyterian Hospital/Weill Cornell Medical Center, New York, NY, USA, 10065

* Corresponding Authors: Elvisha Dhamala (eld2024@med.cornell.edu) and Amy Kuceyeski (amk2012@med.cornell.edu)

Acknowledgements

This work was supported by the following National Institutes of Health (NIH) grants: R21 NS104634-01 (AK) and R01 NS102646-01A1 (AK). The sponsor did not have any role in the study design, the analysis, or interpretation of the data; the writing of the report; or the decision to submit the manuscript for publication.

Data were provided and made available by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centres that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Centre for Systems Neuroscience at Washington University.

Abstract

A thorough understanding of sex-independent and sex-specific neurobiological features that underlie cognitive abilities in healthy individuals is essential for the study of neurological illnesses in which males and females differentially experience and exhibit cognitive impairment. Here, we evaluate sex-independent and sex-specific relationships between functional connectivity and individual cognitive abilities in 392 healthy young adults (196 males) from the Human Connectome Project. First, we establish that sex-independent models comparably predict crystallised abilities in males and females, but more accurately predict fluid abilities in males. Second, we demonstrate sex-specific models comparably predict crystallised abilities within and between sexes, and generally fail to predict fluid abilities in either sex. Third, we reveal that largely overlapping connections between visual, dorsal attention, ventral attention, and default mode networks are associated with better performance on crystallised and fluid cognitive tests in males and females, while connections within visual, somatomotor, and default mode networks are associated with poorer performance. Together, our findings suggest that shared neurobiological features of the functional connectome underlie crystallised and fluid abilities across the sexes.

Keywords: neuroimaging, connectomics, machine learning, functional connectivity, cognition, prediction, sex-specific

Introduction

Cognitive sex differences are widely studied and controversial in neuroscience. Studies often report contradictory sex differences in cognitive abilities, and many are not replicated (Camarata & Woodcock, 2006; Fairweather, 1976; Gur & Gur, 2017; Irwing & Lynn, 2005; Kimura, 2004; Lynn & Irwing, 2004; Satterthwaite et al., 2015; Weis, Hodgetts, & Hausmann, 2019). In recent years, sex differences in cognitive manifestations of various neurological, neurodevelopmental, and neuropsychiatric illnesses have become increasingly evident (Han et al., 2012; Irvine, Laws, Gale, & Kondel, 2012; Laws, Irvine, & Gale, 2016; Subramaniapillai, Almey, Rajah, & Einstein, 2020). Insight into sex-independent and sex-specific brain-behaviour relationships in healthy young adults can enable better understanding of the neurobiological underpinnings of cognitive deficits within and across sexes, paving the way for the development and implementation of personalised treatment strategies. In this study, we aim to disentangle sex-specific and sex-independent brain-behaviour relationships between resting-state functional connectivity and cognitive abilities in healthy young adults.

Resting-state functional connectivity is defined as the temporal dependence of the blood-oxygen-level dependent (BOLD) response in anatomically separate brain regions at rest (Aertsen, Gerstein, Habib, & Palm, 1989; Friston, Frith, Liddle, & Frackowiak, 1993; Martijn P Van Den Heuvel & Pol, 2010). Many studies have linked functional connectivity to cognitive functioning (Casey, Galvan, & Hare, 2005; Casey, Giedd, & Thomas, 2000; Cole, Yarkoni, Repovs, Anticevic, & Braver, 2012; Moeller, Willmes, & Klein, 2015; Park & Friston, 2013; Seeley et al., 2007; Spreng, Stevens, Chamberlain, Gilmore, & Schacter,

2010; M. P. van den Heuvel, Stam, Kahn, & Hulshoff Pol, 2009) and many have predicted individual cognitive abilities from functional connectivity (Chen et al., 2020; Dhamala, Jamison, Jaywant, Dennis, & Kuceyeski, 2020; He et al., 2020; Li et al., 2019; Zimmermann, Griffiths, & McIntosh, 2018). Recent work in this area has shown global signal regression, or removal of trends in the fMRI signal, improves prediction accuracy (Li et al., 2019), machine and deep learning models perform comparably (He et al., 2020), and shared network features predict scores from distinct cognitive domains (Chen et al., 2020; Dhamala et al., 2020). These studies aim to capture brain-behaviour relationships that exist between functional connectivity and cognitive abilities, but it remains unclear whether these relationships are consistent across the sexes.

Sex differences in functional connectivity have been observed across distinct populations (Cummings et al., 2020; Gong, He, & Evans, 2011; Gur & Gur, 2017; Kogler et al., 2016; Satterthwaite et al., 2015; Scheinost et al., 2015; Weis, Hodgetts, et al., 2019; Zhang, Dougherty, Baum, White, & Michael, 2018). Previous work in a developmental cohort has shown males exhibit stronger inter-network connectivity, while females exhibit stronger intra-network connectivity (Satterthwaite et al., 2015). Extant literature also suggests hormonal modulation of functional connectivity (Dubol et al., 2020; Fitzgerald, Pritschet, Santander, Grafton, & Jacobs, 2020; Hjelmervik, Hausmann, Osnes, Westerhausen, & Specht, 2014; Pritschet et al., 2020; Weis, Hodgetts, et al., 2019). In terms of functional connectivity features that discriminate sex, two studies identified that connections within and between frontoparietal and default mode networks strongly contribute to the predictions (Weis, Patil, et al., 2019; Zhang et al., 2018). Together, these studies suggest

sex differences exist in functional organisation of the brain, but whether they map onto cognitive differences remains a mystery.

A recent study similar to this one investigated differences between males and females in predictability of individual intelligence quotient (IQ) and sub-domain cognitive scores using whole-brain functional connectivity (Jiang, Calhoun, Fan, et al., 2020). Their individualized prediction integrated feature selection and regression with a leave-one-out cross validation strategy, resulting in distinct functional connectivity features being selected for each interaction. They reported IQ and other cognitive scores are generally more predictable in females than they are in males, and the sex-specific models rely on distinct functional connections to make predictions. A second study from the same group used a similar approach to predict IQ in males and females using functional connectivity, cortical thickness, or both (Jiang, Calhoun, Cui, et al., 2020). The reported no differences in prediction accuracy between males and females but found that sex-specific models relied on distinct neurobiological correlates. While these findings suggest the presence of distinct brain-behaviour relationships across the sexes, their leave-one-out prediction approach, resulting in distinct features for every iteration, limits the extent to which we can compare and generalise these results because the features used are dependent on which subject is left out in the cross validation. In this current study, we aim to address this concern and expand upon this work.

Here, we study sex-independent and sex-specific brain-behaviour relationships between functional connectivity and individual cognitive abilities in 392 healthy young adults (196 male-female pairs matched for cognitive composite scores) from the Human Connectome Project (Van Essen et al., 2013). First, we quantify whether sex-independent models differ

in how accurately they can predict distinct cognitive abilities from functional connectivity in males and females. Second, we quantify whether sex-specific models better predict individual cognitive abilities from functional connectivity within or between sexes. Third, we evaluate whether shared or sex-specific functional connectivity features map to cognitive abilities.

Methods

The methods used here build upon our prior work (Dhamala et al., 2020) but the analyses presented are novel and aim to identify shared and sex-specific features that predict cognitive abilities. Our experimental workflow is shown in Figure 1. The data that support the findings of this study are openly available as part of the Human Connectome Project at https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release (Van Essen et al., 2013). Codes used to generate the results presented here are available on GitHub (https://github.com/elvisha/SexSpecificCognitivePredictions).

Dataset: We used publicly-available high resolution, preprocessed MRI data from the Human Connectome Project (HCP) – Young Adult S1200 release (Van Essen et al., 2013). MRI data were acquired on a Siemens Skyra 3T scanner at Washington University in St. Louis. Acquisitions included T1-weighted and T2-weighted anatomical images (0.7mm isotropic), and functional MRI (2.0mm isotropic, TR/TE = 720/33.1ms, 8x multiband acceleration). Functional MRI were collected with both left-right and right-left phase encoding. We examined resting-state functional MRI (rfMRI) time series from 196 male-female pairs (n=392) of unrelated healthy young adults with four complete rfMRI runs that were matched for their cognitive composite scores. Although the term *gender* is used in the HCP Data Dictionary, we use the term *sex* in this article because the database collected self-reported biological sex information as opposed to gender identification. We did not verify the self-reported biological sex using genetic information.

Parcellation: We used a subject-specific CoCo439 parcellation that was developed in-house by combining parts of several atlases. This parcellation includes 358 (of 360) functionally derived cortical regions from HCP multi-modal parcellation (MMP) (Matthew

F Glasser et al., 2016) (two hippocampal regions were excluded as they were included in other subcortical ROIs); 12 anatomically defined subcortical regions derived from FreeSurfer's aseg.mgz, adjusted by FSL's FIRST tool (Patenaude, Smith, Kennedy, & Jenkinson, 2011); 12 anatomically defined subcortical nuclei from AAL3v1 (Rolls, Huang, Lin, Feng, & Joliot, 2020); 30 anatomically defined subcortical nuclei from FreeSurfer 7 (Iglesias et al., 2018) (50 nuclei were merged down to 30 to remove the smallest nuclei, as with AAL3v1); and 27 anatomically defined cerebellar regions from the SUIT atlas (Diedrichsen, Balsters, Flavell, Cussans, & Ramnani, 2009). Additional details and corresponding files for this parcellation are available on GitHub (https://github.com/kjamison/nemo#parcellations).

Functional Connectivity Extraction: Each subject underwent four gradient-echo EPI resting-state fMRI (rsfMRI) runs of ~15 min each over two sessions. There are 1200 volumes per scan for a total of 4800 volumes for each subject over the four runs. The minimal preprocessing pipeline performed by the HCP consortium included motion and distortion correction, registration to subject anatomy and standard MNI space, and automated removal of noise artefacts by independent components analysis (M. F. Glasser et al., 2013; Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). We regressed the global signal and its temporal derivative from each rsfMRI time series and concatenated the four scans. We then computed the zero lag Pearson correlation between the concatenated time series from each pair of regions to derive the functional connectivity matrix, which we then Fisher's z-transformed. We used the vectorised upper triangular of this functional connectivity matrix to predict cognition.

Cognition: The NIH Toolbox Cognition Battery is an extensively validated battery of neuropsychological tasks (Carlozzi et al., 2017; Gershon et al., 2013; Heaton et al., 2014; Mungas et al., 2014; Tulsky et al., 2017; Weintraub et al., 2013; Weintraub et al., 2014; Zelazo et al., 2014) that assesses five cognitive domains: language, executive function, episodic memory, processing speed, and working memory through seven individual test instruments (Heaton et al., 2014). The specific tasks include Dimensional Change Card Sort Test, Flanker Inhibitory Control and Attention Test, Picture Sequence Memory Test, Picture Vocabulary Test, Oral Reading Recognition Test, List Sorting Working Memory Test, and Pattern Comparison Processing Speed (Heaton et al., 2014). Three composite scores are derived from participants' scores on the NIH Toolbox Cognitive Battery tasks: Crystallised Cognition Composite, Fluid Cognition Composite, and Total Cognition Composite (Heaton et al., 2014). The Crystallised Cognition Composite comprises the Picture Vocabulary and Oral Reading Recognition tests and assesses language and verbal skills. The Fluid Cognition Composite comprises scores on the Dimensional Change Card Sort, Flanker Inhibitory Control and Attention, Picture Sequence Memory, List Sorting Working Memory, and Pattern Comparison Processing Speed tests. It is a composite that broadly assesses processing speed, memory, and executive functioning. The Total Cognition Composite combines the Crystallised and Fluid Cognition Composites. Composite scores tend to be more reliable/stable but do not capture variability in individual tasks (Heaton et al., 2014). In this study, we investigated the Crystallised, Fluid, and Total Cognition Composites, along with the individual scores from the seven tasks comprising them.

Prediction of Cognitive Performance: We used functional connectivity to predict ten distinct outputs (three composite scores and seven task scores). For each prediction, we trained three distinct models: one sex-independent (trained on both male and female subjects), and two sex-specific (one trained on males, and one trained on females). For each model, we randomly shuffled and split the data into 100 distinct training (80%) and testing (20%) splits. We fit a linear ridge regression model on Scikit-learn (Pedregosa et al., 2011) using the training subset and tuned the regularisation parameter with five iterations of nested cross validation with three-fold inner and outer loops. We optimised the regularisation parameter in the inner loop and validated it in the outer loop. We took the median optimised hyperparameters from the five iterations to generate a single final model. We trained this model on the entire training set, extracted feature weights, and evaluated the model's prediction accuracy on two distinct hold-out test sets: one test set comprised of male subjects and the other comprised of female subjects. We quantify prediction accuracy as the Pearson correlation between the true and predicted values (Li et al., 2019). Across the 100 iterations of each model, we kept the distinct train/test splits consistent for males and females.

Model Significance: For each predictive model, we generated a corresponding null distribution for assessing model significance in the following way. First, for each of the original 100 train/test splits, we permuted both sets' predicted variables (cognitive score) 250 times while preserving the same data within the train/test split. For each of these 25,000 permutations, we trained and tested the model on the permuted data to obtain a null distribution of model performance. We assessed whether the original model's performance was significantly non-zero by comparing the prediction accuracy from each

of the original model's 100 train/test splits to the median prediction accuracy from the null distribution. Specifically, the p-value for the model's significance is the proportion of 100 original models that had prediction accuracies less than or equal to the median performance of the null model. We then corrected the p-values for multiple comparisons over all models (trained on both sexes, males only, and females only to predict ten distinct cognitive scores) and both test subsets (males only and females only) using the Benjamini-Hochberg False Discovery Rate (q=0.05) procedure (Benjamini & Hochberg, 1995).

Model Comparisons: For each cognitive score, our workflow generated two distributions of 100 performance values: the first representing model performance when evaluated on only male individuals, and the second representing model performance when evaluated on only female individuals. For each cognitive score, we compared prediction performance across the male and female test sets using an exact test of differences (MacKinnon, 2009).

Feature Importance**:** We adjusted feature weights to increase their interpretability as described in (Haufe et al., 2014). Briefly, for each iteration of a model, we used the feature weights, W, the covariance of the input variable (functional connectivity) in the training set, Σ_x, and the covariance of the output variable (cognitive score) in the training set, Σ_y, to extract the adjusted feature weights, A, as follows:

$$A = \Sigma_x W \Sigma_y^{-1}$$

We then averaged the adjusted feature weights over the 100 iterations of each model to obtain feature importance matrices. Pairwise regional feature importances were mapped to the network level (Figure S1) by assigning each cortical region from the CoCo439 atlas

to one of 17 networks from the Yeo 17-network parcellation (Yeo et al., 2011). Subcortical regions in the CoCo439 atlas were assigned to a subcortical network, and cerebellar regions to a cerebellar network. The average of the positive and negative feature importances of region pairs within and between the 17 networks were calculated separately; the result is a set of positive and negative importance of connections between and within the 17 networks. We evaluated the Pearson correlation between different models' pairwise network-level feature importances, where positive and negative importances were considered together by concatenating them into a single vector.

Figure 1: Experimental workflow. a) First, we generated individual functional connectivity using Pearson correlation of regional global signal regressed resting-state functional MRI time series. b) Second, we compiled cognitive scores for all subjects. The NIH Toolbox Cognition Battery assesses five cognitive domains using seven tests. The Crystallised Cognition Composite (blue) reflects language (vocabulary, reading decoding). The Fluid Cognition Composite (green) reflects executive function (cognitive flexibility, inhibitory control and attention), episodic memory, working memory, and processing speed. The Total Cognition Composite (dotted) combines the Crystallised and Fluid Composite scores. c) Third, we predicted each cognitive score from functional connectivity using sex-independent and sex-specific linear ridge regression models. We randomly shuffled and split the male and female subjects into train (80%) and test (20%) groups. Male and female training subsets were concatenated for the sex-independent models and kept separate for the sex-specific models. We performed five shuffled iterations of nested cross validation with three-fold inner and outer loops. The model hyperparameter was optimised in the inner loop and validated in the outer loop. The median optimised hyperparameter from five iterations of nested cross validation was used to train the final model on the entire (sex-independent or sex-specific) training set and evaluated on the (sex-independent or sex-specific) test hold-out set. This was repeated for 100 unique train/test splits.

[Insert Figure 1 here]

Results

Sex-Independent Models: Sex-independent models significantly predict Total and Crystallised Composite scores for both sexes, and Fluid Composite scores in males only, corrected $p<0.05$ (Figure 2). Within the crystallised domain, we significantly predict Picture Vocabulary scores in both sexes, but only significantly predict Reading scores in females. Within the fluid domain, we significantly predict Dimensional Change Card Sort, Picture Sequence Memory, and Processing Speed scores in males, while we fail to significantly predict Flanker and List Sorting scores in males or females.

Sex-Specific Models: Sex-specific male-trained and female-trained models significantly predict Total Composite scores in both sexes (Figure 3). Using female-trained models, we significantly predict Crystallised Composite scores in both sexes but fail to significantly predict Fluid Composite scores in either sex. Using male-trained models, we significantly predict Crystallised Composite scores in females and Fluid Composite scores in males. Within the crystallised domain, we significantly predict Picture Vocabulary scores in both sexes using both sex-specific models, but only significantly predict Reading scores in the opposite sex. Within the fluid domain, we significantly predict Dimensional Change Card Sort in males using male-trained models but fail to significantly predict all Flanker, Picture Sequence Memory, List Sorting, and Processing Speed scores in either sex using either sex-specific model.

Figure 2: Violin plots of accuracy for sex-independent models in predicting cognitive composite scores and individual task scores. Blue violins represent accuracy of models tested on male subjects and red represents of models tested on female subjects. The shape of the violin plots indicates the entire distribution of values, dashed lines indicate the median, and dotted lines indicate the interquartile range. Solid colour violin plots represent models that performed above chance levels based on permutation tests. Vertical dotted lines separate individual tests according to cognitive domain: general, crystallised, and fluid.

[Insert Figure 2 here]

Figure 3: Violin plots of accuracy for sex-specific models in predicting cognitive composite scores and individual task scores. Purple indicates results from models trained and tested on males; blue indicates results from models trained on males and tested on females; green indicates results from models trained on females and tested on males; and orange indicates results from models trained and tested on females. The shape of the violin plots indicates the entire distribution of values, dashed lines indicate the median, and dotted lines indicate the interquartile range. Solid colour violin plots indicate those models that performed above chance levels based on permutation tests. Vertical dotted lines separate individual tests according to cognitive domain: general, crystallised, and fluid.

[Insert Figure 3 here]

Model Comparisons: Using an exact test of differences, we did not identify any significant differences in model performance between the sexes in the sex-independent models or the sex-specific models for any cognitive score.

Feature Importance Comparisons: We correlated network-level feature importances between the sex-independent and sex-specific models (Figure 4). Features important in predicting the Total Composite, Crystallised Composite, and specific crystallised task scores from the sex-independent models are equally correlated to those from the male- and female- specific models. Features important in predicting the Fluid Composite and specific fluid scores are more strongly correlated to features important to predict those scores in males than in females. Features important in predicting each of the scores from the sex-specific models are generally more strongly correlated within sexes for different cognitive scores than across sexes for the same cognitive score; however, the correlations between models trained on different sexes is generally high. Features important in predicting the Total Composite score are correlated with features important to predict the Crystallised and Fluid composite scores and each of the individual task scores. Feature importance for predicting specific crystallised task scores are more strongly correlated with feature importance for predicting the Crystallised Composite score in females than they are in males. Features important for predicting specific fluid task scores are more strongly correlated to those important for predicting the Fluid Composite score in males than they are in females.

Figure 4: Pearson correlation of network-level feature importance for the sex-independent and sex-specific models predicting each cognitive score. Positive and negative network-level feature importance were computed by taking the positive and negative sums of the regional feature importance. Correlations were evaluated between the concatenated positive and negative network-level feature importances.

[Insert Figure 4 here]

Network-Level Feature Importance: Stronger functional connections between visual, dorsal/ventral attention, and default mode networks are associated with higher crystallised abilities in males and females (Figure 5). Stronger functional connections within and between visual, temporal parietal, dorsal attention, default mode, and limbic networks predict higher fluid abilities in females, while stronger functional connections between visual and dorsal attention networks predict higher fluid abilities in males. Stronger functional connections within visual, somatomotor, and default mode networks predict lower fluid and crystallized abilities in both sexes. Generally similar functional connections predict Picture Vocabulary and Reading scores in both sexes (Figure 6) as well as scores in individual fluid tasks, with the exception of List Sort and Picture Sequence scores (Figure S2). In females, stronger functional connections within visual, temporal parietal, dorsal attention, control, default, and limbic networks predict higher List Sort scores, while stronger connections between those networks predict lower scores. In males, stronger connections between visual and dorsal attention, and within dorsal attention, limbic, and control network predict higher List Sort scores, while stronger connections within visual, somatomotor, and default mode network predict lower scores. Stronger functional connections within visual and default mode networks predict higher Picture Sequence scores in females, while stronger connections within limbic networks, and between visual, dorsal attention, and default mode networks predict higher scores in males.

Figure 5: Network-level positive and negative feature importance for females (left two columns) and males (right two columns) to predict crystallised (top), fluid (middle), and total (bottom) cognition composites. Node radii and colour denote strength of intra-network feature importance. Edge thickness and colour denote strength of inter-network feature importance. Warmer colours are used for positive feature importance, and cooler colours for negative feature importance.

[Insert Figure 5 here]

Figure 6: Network-level positive and negative feature importance for females (left two columns) and males (right two columns) to predict individual crystallised cognition task scores: picture vocabulary (top) and reading (bottom). Node radii and colour denote strength of intra-network feature importance. Edge weight and colour denote strength of inter-network feature importance. Warmer colours are used for positive feature importance, and cooler colours for negative feature importance.

[Insert Figure 6 here]

Discussion

In this study, we quantified sex-independent and sex-specific relationships between functional connectivity and cognition. Using whole brain resting-state functional connectivity, we predicted individual crystallised and fluid abilities in 392 healthy young adults. First, we find sex-independent models predict with equivalent accuracy crystallised abilities in both sexes but predict fluid abilities more accurately in males. Second, we show sex-specific models perform comparably when predicting crystallised abilities within and between sexes, but generally fail to predict fluid abilities in either sex, except for the fluid composite and Dimensional Change Card Sort score in males. Third, we demonstrate that sex-specific models predicting crystallised and fluid abilities generally rely on shared functional connections within and between distinct cortical networks. Together, our findings largely suggest that shared neurobiological features predict general and specific crystallised abilities in both sexes.

Crystallised cognition primarily represents language (vocabulary and reading decoding) abilities, while fluid cognition represents a wider range of cognitive processes including executive function (cognitive flexibility and inhibitory control and attention), episodic memory, working memory, and processing speed. Prior work has shown Total and Crystallised Composite scores are more predictable than the Fluid Composite (Dhamala et al., 2020) but that work did not investigate whether the same is true for specific tasks within the cognitive domains or whether these results hold equally among males and females. In this current work, we replicate and expand upon those previous findings.

In this study, we used permutation tests to evaluate whether models are capturing signal or noise: performance above chance levels suggests the model is representing

meaningful brain-behaviour relationships while performance below chance levels suggests it is primarily representing noise. We also used exact tests of differences to compare whether the models are significantly different from one another in terms of their performance. Although our exact tests for differences showed a lack of performance differences between the sexes in all sex-independent and sex-specific models, our permutation tests indicate otherwise. Even though a pair of models may not be significantly different from one another in their performance (based on the exact test of differences), if one is performing above chance levels and the other is not, it suggests one model is capturing meaningful information while the other is not. Results from our sex-independent models suggest they might be capturing shared relationships between functional connectivity and crystallised abilities in males and females, but male-specific relationships between functional connectivity and fluid abilities. This is supported by our observation that connectivity-cognition relationships for fluid abilities from the sex-independent models more closely resemble those from the male-specific models than the female-specific models. Results from our sex-specific models provide additional support for our findings from the sex-independent models, as we find that connectivity-cognition relationships for crystallised abilities and overall cognition are generally shared between the sexes. We also observe an even greater inability to predict fluid abilities with our sex-specific models compared to our sex-independent models, which could be in part due to the decreased sample size in the sex-specific models. The general lack of predictability observed for fluid abilities in both types of models may be underscored by individual differences in the signal-to-noise ratio of the specific brain-behaviour relationships. Fluid abilities are more susceptible to factors including sleep, stress, and mood which directly

influence executive functions and memory and less stable within an individual over time (Nilsson et al., 2005; O'Neill, Kamper-DeMarco, Chen, & Orom, 2020; Salthouse, 2010). Our inability to accurately predict most fluid abilities with our models provides support for the null hypothesis that fluid abilities/executive function are not strongly related to functional connectivity. Hence, we primarily focus our interpretation on results pertaining to the crystallised cognitive domain.

Our understanding of cognitive sex differences and brain-behaviour relationships have widely shifted over the decades. While research has confirmed some differences, many others have been refuted (Halpern, 2013; Miller & Halpern, 2014). Two similar studies to date investigating sex-specific brain-behaviour relationships have reported contradictory findings. Implementing a connectome-based prediction modelling approach, Jiang et al observed no differences in prediction accuracy between males and females when predicting IQ using functional connectivity (Jiang, Calhoun, Cui, et al., 2020). In a second study from the group, they demonstrated IQ was more predictable in females than in males (Jiang, Calhoun, Fan, et al., 2020). In this current work, our sex-independent models comparably predict overall cognition and crystallised abilities in males and females, but better predict some fluid abilities in males compared to females while failing to predict other fluid tasks in either sex altogether. In this study, we implemented a nested cross validation approach with 100 different randomised splits of the data to generate a distribution of performance accuracy measures. Previous studies relied on integrating feature selection with a leave-one-out cross validation approach resulting in a single accuracy value for the model and distinct features being used to predict the output variable for each subject (Jiang, Calhoun, Cui, et al., 2020; Jiang, Calhoun, Fan, et al.,

2020). Due to these methodological differences, our prediction accuracy results cannot be directly compared to prior work. However, it is worth noting that our sex-specific models comparably predict overall and crystallised aspects of cognition in males and females, supporting one of the previous studies (Jiang, Calhoun, Cui, et al., 2020) but contradicting the other (Jiang, Calhoun, Fan, et al., 2020).

In this study, we find connections within and between distinct cortical networks are crucial to predict cognition, and these features are shared between the sexes, contradicting extant literature implementing sex-specific models (Jiang, Calhoun, Cui, et al., 2020; Jiang, Calhoun, Fan, et al., 2020). More specifically, we find stronger connections between the visual, dorsal attention, ventral attention, and default mode networks predict higher crystallised and fluid ability scores in both sexes, while stronger connections within visual, somatomotor, and default mode networks predict lower crystallised and fluid ability scores in both sexes. While some differences in male and female models' feature importances exist, their correlations are moderate to high (R = 0.6-0.9). One notable difference between the sexes is in the positive feature importance map for the fluid composite. However, the female model did not perform better than chance for fluid ability predictions while the male model did, limiting the relevance of this finding. We also demonstrate that feature importance correlations, within and between sexes, are stronger for tasks within the crystallised domain than tasks within the fluid domain or tasks between the two domains. This is likely related to the models' overall lower accuracies in predicting fluid abilities; if the models are not reliably mapping functional connectivity to fluid abilities, there will be more noise in their feature importance, resulting in lower correlations across models. Our results contradict findings from prior work identifying distinct correlates of

cognition in males and females. In one study, authors reported the top 100 functional connections to predict IQ in males and females are distinct with only three overlapping features (Jiang, Calhoun, Fan, et al., 2020). In a second study, authors found male IQ was more strongly correlated with functional connectivity in left parahippocampus and default mode network, while female IQ was more strongly correlated with functional connectivity in putamen and cerebellar network (Jiang, Calhoun, Cui, et al., 2020). This discrepancy in findings could be due to model differences, particularly in the cross-validation, feature selection and inference choices, or the choice of cognitive score.

Limitations:

In this study, we trained and tested sex-independent and sex-specific models on 196 male and 196 female subjects, all unrelated. Over each of the 100 outer loop permutations, we ensured the same set of male/female subjects were in the training and testing subsets for the sex-independent and the male/female-specific models. Maintaining this consistency of subjects allowed us to maintain the variance within the subjects, but also resulted in our sex-independent models being trained and tested on twice as many subjects as our sex-specific models. Prior work has demonstrated that fluid abilities are more difficult to predict than crystallised abilities (Dhamala et al., 2020). In this study, we found sex-independent models were able to predict some fluid abilities above chance levels in males, but sex-specific models generally did not perform above chance levels for either sex. The inherent difficulty in predicting fluid abilities, combined with the lower number of subjects for the sex-specific models, may explain why many of our sex-specific models performed poorly. In this study, our main goal was to evaluate whether the models differed in their predictions of cognitive abilities between males and females rather than between

the models themselves. However, future work in this area should explore whether sex-independent and sex-specific models differ from one another when training sample sizes are consistent.

Many researchers studying cognitive differences between males and females compare group averages between the sexes. While this approach can yield insightful results pertaining to general sex differences, their relevance to individual cognitive abilities in males and females is limited. Genetic, hormonal, cultural, and psychosocial factors can influence sex-related and sex-independent individual differences in functional connectivity and cognition (Cosgrove, Mazure, & Staley, 2007; Miller & Halpern, 2014). Here, we sought to uncover whether relationships between functional connectivity and cognition are shared between the sexes or are distinct. Our results largely suggest shared network connectivity features equally predict cognitive abilities in males and females. However, we must acknowledge that here, due to the limitations of the data set, we can only consider individuals' sex but not their gender identity or fluidity. Our society projects distinct gender roles onto males and females paving the way for a lifetime of gender-differentiated experiences (Eliot, 2011). These distinct social factors may drive gender differences in brain-behaviour relationships, even in the absence of sex differences, that our study is not designed to capture. Future work in this area should aim to collect and integrate data about gender identity and fluidity so we can better understand how relationships between connectivity and cognition may or may not vary with gender.

Many machine learning models based on neuroimaging data struggle with generalisability due to differences in study sites, scanner types, and scan parameters. The models we have designed in this study were only trained, validated, and tested on data from the

Human Connectome Project. Although we implement a nested cross validation approach and evaluate our models with 100 distinct train/test splits, the results we report may not be entirely comparable or generalisable to other datasets. Future studies should aim to integrate data from multiple sites to address this limitation.

Conclusion

A comprehensive understanding of neurobiological markers that underlie cognitive abilities within and across sexes is necessary if we are to understand sex-specific effects of aging and illness on cognition. Here, we implement predictive modelling approaches to explore sex-independent and sex-specific relationships between functional connectivity and cognitive abilities. We report three main findings. We demonstrate that sex-independent models comparably capture relationships between connectivity and crystallised abilities in males and females, but only successfully capture relationships between connectivity and fluid abilities in males. We find sex-specific models comparably predict crystallised abilities within and between sexes, but fail to predict fluid abilities in either sex. Finally, we find that stronger connections between visual, dorsal attention, ventral attention, and default mode networks predict higher crystallised and fluid ability scores, and stronger connections within visual, somatomotor, and default mode networks predict lower crystallised and fluid ability scores in both sexes. Taken together, this suggests that brain-behaviour relationships are shared between the sexes and rely on overlapping network connectivity within and between cortical structures.

Citation Gender Diversity Statement: Recent work in neuroscience and other fields has identified a bias in citation practices such that papers from women and other minorities are under-cited relative to the number of such papers in the field (Caplar, Tacchella, & Birrer, 2017; Chakravartty, Kuo, Grubbs, & McIlwain, 2018; Dion, Sumner, & Mitchell, 2018; Dworkin et al., 2020; Maliniak, Powers, & Walter, 2013; Thiem, Sealey, Ferrer, Trott, & Kennison, 2018). Here we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, and other factors. We used classification of gender based on the first names of the first and last authors (Dworkin et al., 2020), with possible combinations including male/male, male/female, female/male, and female/female. Excluding self-citations to the first and last authors of our current paper, the references contain 50.0% male/male, 17.6% male/female, 16.2% female/male, and 16.2% female/female. We look forward to future work that could help us to better understand how to support equitable practices in science.

Author Contributions: Conceptualization, E.D., A.K.; Methodology, E.D., K.W.J., A.K.; Software, E.D.; Investigation, E.D., A.J., A.K.; Formal Analysis, E.D., A.K.; Resources, A.K; Data Curation, E.D., K.W.J.; Writing – Original Draft, E.D.; Writing – Review & Editing, E.D., K.W.J., A.J., A.K.; Visualisation, E.D.; Supervision, A.K.; Funding Acquisition, A.K.

Data Availability Statement: The data used are openly available as part of the Human Connectome Project at https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release.

Conflict of Interest Statement: The authors declare no competing financial interests.

References

Aertsen, A., Gerstein, G., Habib, M., & Palm, G. (1989). Dynamics of neuronal firing correlation: Modulation of" effective connectivity". *Journal of neurophysiology, 61*(5), 900-917.

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-Statistical Methodology, 57*(1), 289-300.

Camarata, S., & Woodcock, R. (2006). Sex differences in processing speed: Developmental effects in males and females. *Intelligence, 34*(3), 231-252.

Caplar, N., Tacchella, S., & Birrer, S. (2017). Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nature Astronomy, 1*(6), 1-5.

Carlozzi, N. E., Tulsky, D. S., Wolf, T. J., Goodnight, S., Heaton, R. K., Casaletto, K. B., . . . Heinemann, A. W. (2017). Construct validity of the nih toolbox cognition battery in individuals with stroke. *Rehabil Psychol, 62*(4), 443-454. doi:10.1037/rep0000195

Casey, B. J., Galvan, A., & Hare, T. A. (2005). Changes in cerebral functional organization during cognitive development. *Curr Opin Neurobiol, 15*(2), 239-244. doi:10.1016/j.conb.2005.03.012

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biol Psychol, 54*(1-3), 241-257.

Chakravartty, P., Kuo, R., Grubbs, V., & McIlwain, C. (2018). # communicationsowhite. *Journal of Communication, 68*(2), 254-266.

Chen, J., Tam, A., Kebets, V., Orban, C., Ooi, L. Q. R., Marek, S., . . . Holmes, A. J. (2020). Shared and unique brain network features predict cognition, personality and mental health in childhood. *bioRxiv*.

Cole, M. W., Yarkoni, T., Repovs, G., Anticevic, A., & Braver, T. S. (2012). Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *J Neurosci, 32*(26), 8988-8999. doi:10.1523/JNEUROSCI.0536-12.2012

Cosgrove, K. P., Mazure, C. M., & Staley, J. K. (2007). Evolving knowledge of sex differences in brain structure, function, and chemistry. *Biol Psychiatry, 62*(8), 847-855. doi:10.1016/j.biopsych.2007.03.001

Cummings, K. K., Lawrence, K. E., Hernandez, L. M., Wood, E. T., Bookheimer, S. Y., Dapretto, M., & Green, S. A. (2020). Sex differences in salience network connectivity and its relationship to sensory over-responsivity in youth with autism spectrum disorder. *Autism Research, 13*(9), 1489-1500.

Dhamala, E., Jamison, K., Jaywant, A., Dennis, S., & Kuceyeski, A. (2020). Integrating multimodal connectivity improves prediction of individual cognitive abilities. *bioRxiv*.

Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., & Ramnani, N. (2009). A probabilistic mr atlas of the human cerebellum. *Neuroimage, 46*(1), 39-46.

Dion, M. L., Sumner, J. L., & Mitchell, S. M. (2018). Gendered citation patterns across political science and social science methodology fields. *Political Analysis, 26*(3), 312-327.

Dubol, M., Neill Epperson, C., Sacher, J., Pletzer, B., Derntl, B., Lanzenberger, R., . . . Comasco, E. (2020). Neuroimaging the menstrual cycle: A multimodal systematic review. *Frontiers in Neuroendocrinology*.

Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. *Nat Neurosci*. doi:10.1038/s41593-020-0658-y

Eliot, L. (2011). The trouble with sex differences. *Neuron, 72*(6), 895-898. doi:10.1016/j.neuron.2011.12.001

Fairweather, H. (1976). Sex differences in cognition. *Cognition, 4*(3), 231-280.

Fitzgerald, M., Pritschet, L., Santander, T., Grafton, S. T., & Jacobs, E. G. (2020). Cerebellar network organization across the human menstrual cycle. *Scientific reports, 10*(1), 1-11.

Friston, K., Frith, C., Liddle, P., & Frackowiak, R. (1993). Functional connectivity: The principal-component analysis of large (pet) data sets. *Journal of Cerebral Blood Flow & Metabolism, 13*(1), 5-14.

Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). Nih toolbox for assessment of neurological and behavioral function. *Neurology, 80*(11 Suppl 3), S2-6. doi:10.1212/WNL.0b013e3182872e5f

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., . . . Jenkinson, M. (2016). A multi-modal parcellation of human cerebral cortex. *Nature, 536*(7615), 171-178.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., . . . Consortium, W. U.-M. H. (2013). The minimal preprocessing pipelines for

the human connectome project. *Neuroimage, 80*, 105-124. doi:10.1016/j.neuroimage.2013.04.127

Gong, G., He, Y., & Evans, A. C. (2011). Brain connectivity: Gender makes a difference. *Neuroscientist, 17*(5), 575-591. doi:10.1177/1073858410386492

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C. F., Auerbach, E. J., Douaud, G., Sexton, C. E., . . . Smith, S. M. (2014). Ica-based artefact removal and accelerated fmri acquisition for improved resting state network imaging. *Neuroimage, 95*, 232-247. doi:10.1016/j.neuroimage.2014.03.034

Gur, R. C., & Gur, R. E. (2017). Complementarity of sex differences in brain and behavior: From laterality to multimodal neuroimaging. *J Neurosci Res, 95*(1-2), 189-199. doi:10.1002/jnr.23830

Halpern, D. F. (2013). *Sex differences in cognitive abilities*: Psychology press.

Han, M., Huang, X.-F., Xiu, M. H., Hui, L., Liu, H., Kosten, T. R., & Zhang, X. Y. (2012). Gender differences in cognitive function of patients with chronic schizophrenia. *Progress in neuro-psychopharmacology and biological psychiatry, 39*(2), 358-363.

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage, 87*, 96-110.

He, T., Kong, R., Holmes, A. J., Nguyen, M., Sabuncu, M. R., Eickhoff, S. B., . . . Yeo, B. T. (2020). Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *Neuroimage, 206*, 116276.

Heaton, R. K., Akshoomoff, N., Tulsky, D., Mungas, D., Weintraub, S., Dikmen, S., . . . Gershon, R. (2014). Reliability and validity of composite scores from the nih toolbox cognition battery in adults. *J Int Neuropsychol Soc, 20*(6), 588-598. doi:10.1017/S1355617714000241

Hjelmervik, H., Hausmann, M., Osnes, B., Westerhausen, R., & Specht, K. (2014). Resting states are resting traits - an fmri study of sex differences and menstrual cycle effects in resting state cognitive control networks. *PloS one, 9*(7). doi:ARTN e103492

10.1371/journal.pone.0103492

Iglesias, J. E., Insausti, R., Lerma-Usabiaga, G., Bocchetta, M., Van Leemput, K., Greve, D. N., . . . Paz-Alonso, P. M. (2018). A probabilistic atlas of the human thalamic nuclei combining ex vivo mri and histology. *Neuroimage, 183*, 314-326.

Irvine, K., Laws, K. R., Gale, T. M., & Kondel, T. K. (2012). Greater cognitive deterioration in women than men with alzheimer's disease: A meta analysis. *Journal of clinical and experimental neuropsychology, 34*(9), 989-998.

Irwing, P., & Lynn, R. (2005). Sex differences in means and variability on the progressive matrices in university students: A meta-analysis. *British Journal of Psychology, 96*(4), 505-524.

Jiang, R., Calhoun, V. D., Cui, Y., Qi, S., Zhuo, C., Li, J., . . . Jiang, T. (2020). Multimodal data revealed different neurobiological correlates of intelligence between males and females. *Brain imaging and behavior, 14*(5), 1979-1993.

Jiang, R., Calhoun, V. D., Fan, L., Zuo, N., Jung, R., Qi, S., . . . Sui, J. (2020). Gender differences in connectome-based predictions of individualized intelligence quotient and sub-domain scores. *Cerebral cortex, 30*(3), 888-900.

Kimura, D. (2004). Human sex differences in cognition, fact, not predicament. *Sexualities, Evolution & Gender, 6*(1), 45-53.

Kogler, L., Müller, V. I., Seidel, E.-M., Boubela, R., Kalcher, K., Moser, E., . . . Derntl, B. (2016). Sex differences in the functional connectivity of the amygdalae in association with cortisol. *Neuroimage, 134*, 410-423.

Laws, K. R., Irvine, K., & Gale, T. M. (2016). Sex differences in cognitive impairment in alzheimer's disease. *World journal of psychiatry, 6*(1), 54.

Li, J. W., Kong, R., Liegeois, R., Orban, C., Tan, Y. R., Sun, N. B., . . . Yeo, B. T. T. (2019). Global signal regression strengthens association between resting-state functional connectivity and behavior. *Neuroimage, 196*, 126-141. doi:10.1016/j.neuroimage.2019.04.016

Lynn, R., & Irwing, P. (2004). Sex differences on the progressive matrices: A meta-analysis. *Intelligence, 32*(5), 481-498.

MacKinnon, J. G. (2009). Bootstrap hypothesis testing. *Handbook of computational econometrics, 183*, 213.

Maliniak, D., Powers, R., & Walter, B. F. (2013). The gender citation gap in international relations. *International Organization, 67*(4), 889-922.

Miller, D. I., & Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends in cognitive sciences, 18*(1), 37-45.

Moeller, K., Willmes, K., & Klein, E. (2015). A review on functional and structural brain connectivity in numerical cognition. *Front Hum Neurosci, 9*, 227. doi:10.3389/fnhum.2015.00227

Mungas, D., Heaton, R., Tulsky, D., Zelazo, P. D., Slotkin, J., Blitz, D., . . . Gershon, R. (2014). Factor structure, convergent validity, and discriminant validity of the nih toolbox cognitive health battery (nihtb-chb) in adults. *J Int Neuropsychol Soc, 20*(6), 579-587. doi:10.1017/S1355617714000307

Nilsson, J. P., Söderström, M., Karlsson, A. U., Lekander, M., Åkerstedt, T., Lindroth, N. E., & Axelsson, J. (2005). Less effective executive functioning after one night's sleep deprivation. *Journal of sleep research, 14*(1), 1-6.

O'Neill, J., Kamper-DeMarco, K., Chen, X., & Orom, H. (2020). Too stressed to self-regulate? Associations between stress, self-reported executive function, disinhibited eating, and bmi in women. *Eating Behaviors, 39*, 101417.

Park, H. J., & Friston, K. (2013). Structural and functional brain networks: From connections to cognition. *Science, 342*(6158), 1238411. doi:10.1126/science.1238411

Patenaude, B., Smith, S. M., Kennedy, D. N., & Jenkinson, M. (2011). A bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage, 56*(3), 907-922.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, E. (2011). Scikit-learn: Machine learning in python. *Journal of machine learning research, 12*, 2825-2830.

Pritschet, L., Santander, T., Taylor, C. M., Layher, E., Yu, S., Miller, M. B., . . . Jacobs, E. G. (2020). Functional reorganization of brain networks across the human menstrual cycle. *Neuroimage, 220*, 117091.

Rolls, E. T., Huang, C.-C., Lin, C.-P., Feng, J., & Joliot, M. (2020). Automated anatomical labelling atlas 3. *Neuroimage, 206*, 116189.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional mri data: Combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage, 90*, 449-468. doi:10.1016/j.neuroimage.2013.11.046

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society: JINS, 16*(5), 754.

Satterthwaite, T. D., Wolf, D. H., Roalf, D. R., Ruparel, K., Erus, G., Vandekar, S., . . . Gur, R. C. (2015). Linked sex differences in cognition and functional connectivity in youth. *Cereb Cortex, 25*(9), 2383-2394. doi:10.1093/cercor/bhu036

Scheinost, D., Finn, E. S., Tokoglu, F., Shen, X., Papademetris, X., Hampson, M., & Constable, R. T. (2015). Sex differences in normal age trajectories of functional brain networks. *Human brain mapping, 36*(4), 1524-1535.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., . . . Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci, 27*(9), 2349-2356. doi:10.1523/JNEUROSCI.5587-06.2007

Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., & Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *Neuroimage, 53*(1), 303-317.

Subramaniapillai, S., Almey, A., Rajah, M. N., & Einstein, G. (2020). Sex and gender differences in cognitive and brain reserve: Implications for alzheimer's disease in women. *Frontiers in Neuroendocrinology*, 100879.

Thiem, Y., Sealey, K. F., Ferrer, A. E., Trott, A. M., & Kennison, R. (2018). *Just ideas? The status and future of publication ethics in philosophy: A white paper*. Retrieved from

Tulsky, D. S., Holdnack, J. A., Cohen, M. L., Heaton, R. K., Carlozzi, N. E., Wong, A. W. K., . . . Heinemann, A. W. (2017). Factor structure of the nih toolbox cognition battery in individuals with acquired brain injury. *Rehabil Psychol, 62*(4), 435-442. doi:10.1037/rep0000183

Van Den Heuvel, M. P., & Pol, H. E. H. (2010). Exploring the brain network: A review on resting-state fmri functional connectivity. *European neuropsychopharmacology, 20*(8), 519-534.

van den Heuvel, M. P., Stam, C. J., Kahn, R. S., & Hulshoff Pol, H. E. (2009). Efficiency of functional brain networks and intellectual performance. *J Neurosci, 29*(23), 7619-7624. doi:10.1523/JNEUROSCI.1443-09.2009

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & Consortium, W. U.-M. H. (2013). The wu-minn human connectome project: An overview. *Neuroimage, 80*, 62-79. doi:10.1016/j.neuroimage.2013.05.041

Weintraub, S., Bauer, P. J., Zelazo, P. D., Wallner-Allen, K., Dikmen, S. S., Heaton, R. K., . . . Gershon, R. C. (2013). I. Nih toolbox cognition battery (cb): Introduction and pediatric data. *Monogr Soc Res Child Dev, 78*(4), 1-15. doi:10.1111/mono.12031

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Slotkin, J., . . . Gershon, R. (2014). The cognition battery of the nih toolbox for assessment of neurological and behavioral function: Validation in an adult sample. *J Int Neuropsychol Soc, 20*(6), 567-578. doi:10.1017/S1355617714000320

Weis, S., Hodgetts, S., & Hausmann, M. (2019). Sex differences and menstrual cycle effects in cognitive and sensory resting state networks. *Brain and Cognition, 131*, 66-73. doi:10.1016/j.bandc.2017.09.003

Weis, S., Patil, K. R., Hoffstaedter, F., Nostro, A., Yeo, B. T. T., & Eickhoff, S. B. (2019). Sex classification by resting state brain connectivity. *Cereb Cortex*. doi:10.1093/cercor/bhz129

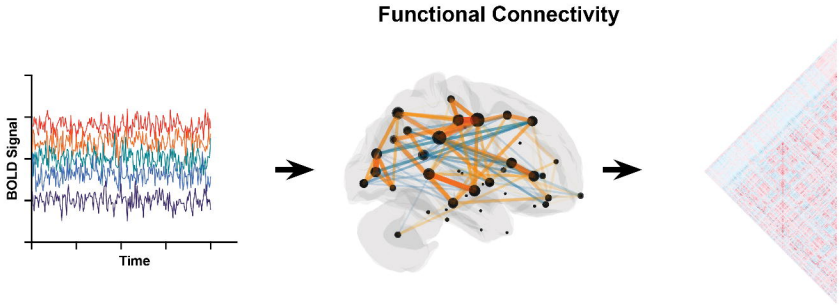
Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., . . . Polimeni, J. R. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology, 106*(3), 1125-1165.

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., Conway, K. P., . . . Weintraub, S. (2014). Nih toolbox cognition battery (cb): Validation of executive function measures in adults. *J Int Neuropsychol Soc, 20*(6), 620-629. doi:10.1017/S1355617714000472

Zhang, C., Dougherty, C. C., Baum, S. A., White, T., & Michael, A. M. (2018). Functional connectivity predicts gender: Evidence for gender differences in resting brain connectivity. *Hum Brain Mapp, 39*(4), 1765-1776. doi:10.1002/hbm.23950

Zimmermann, J., Griffiths, J. D., & McIntosh, A. R. (2018). Unique mapping of structural and functional connectivity on cognition. *J Neurosci, 38*(45), 9658-9667. doi:10.1523/JNEUROSCI.0900-18.2018

a)

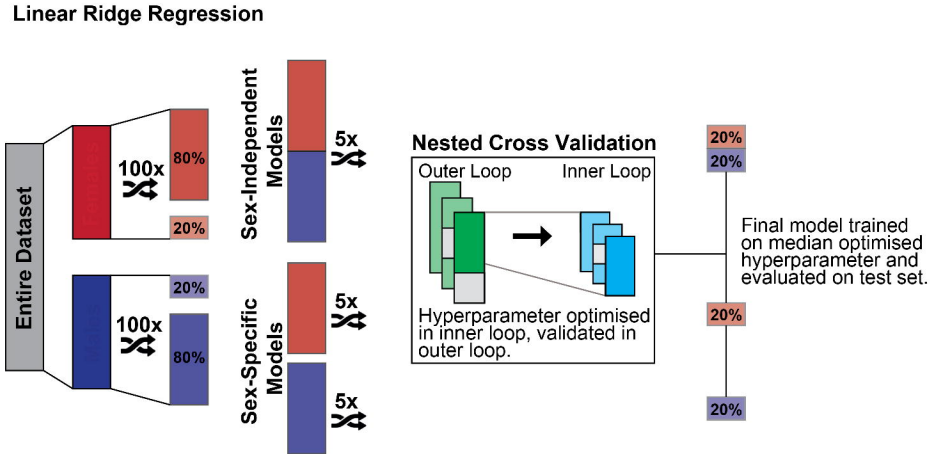


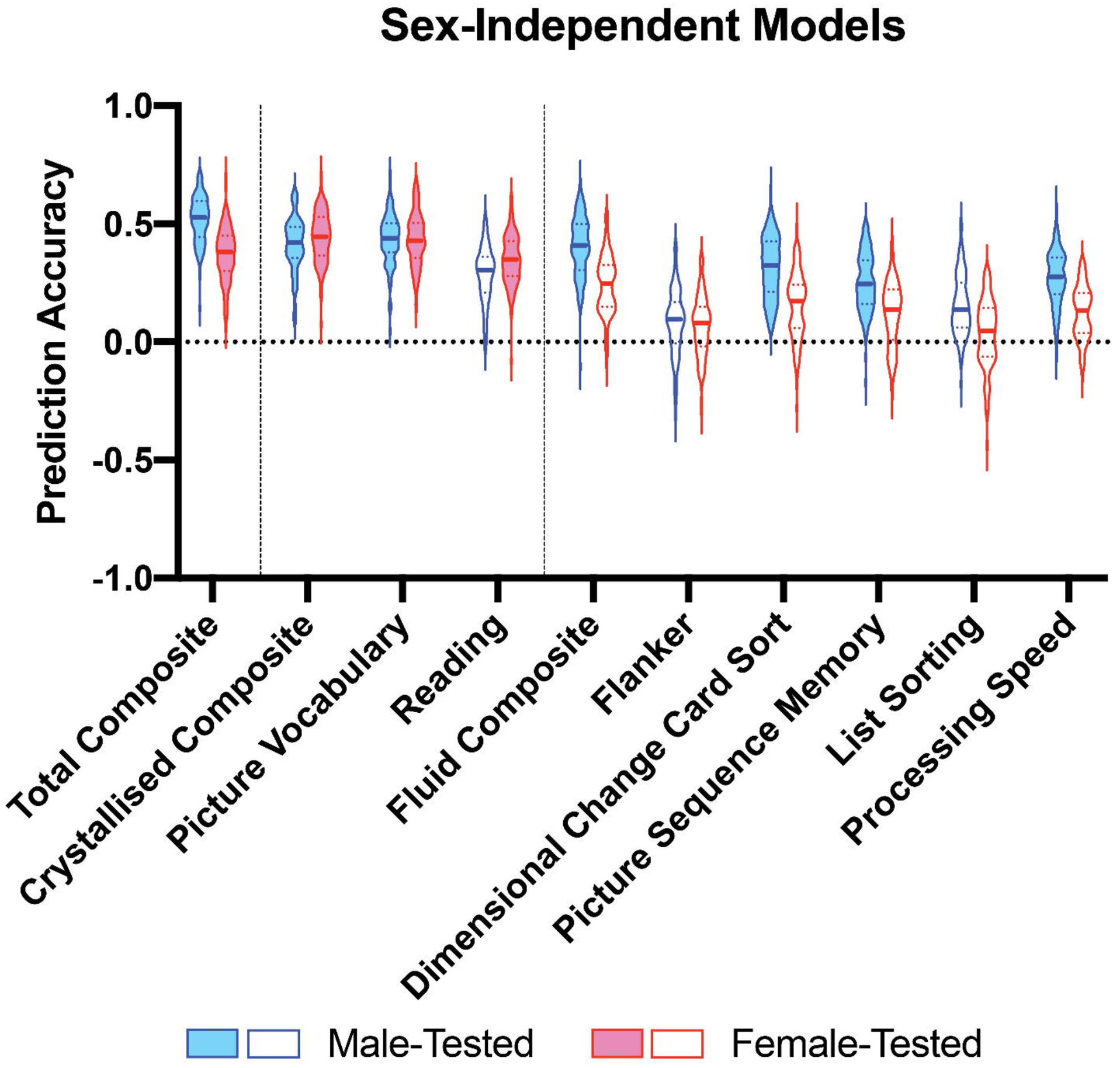
b)

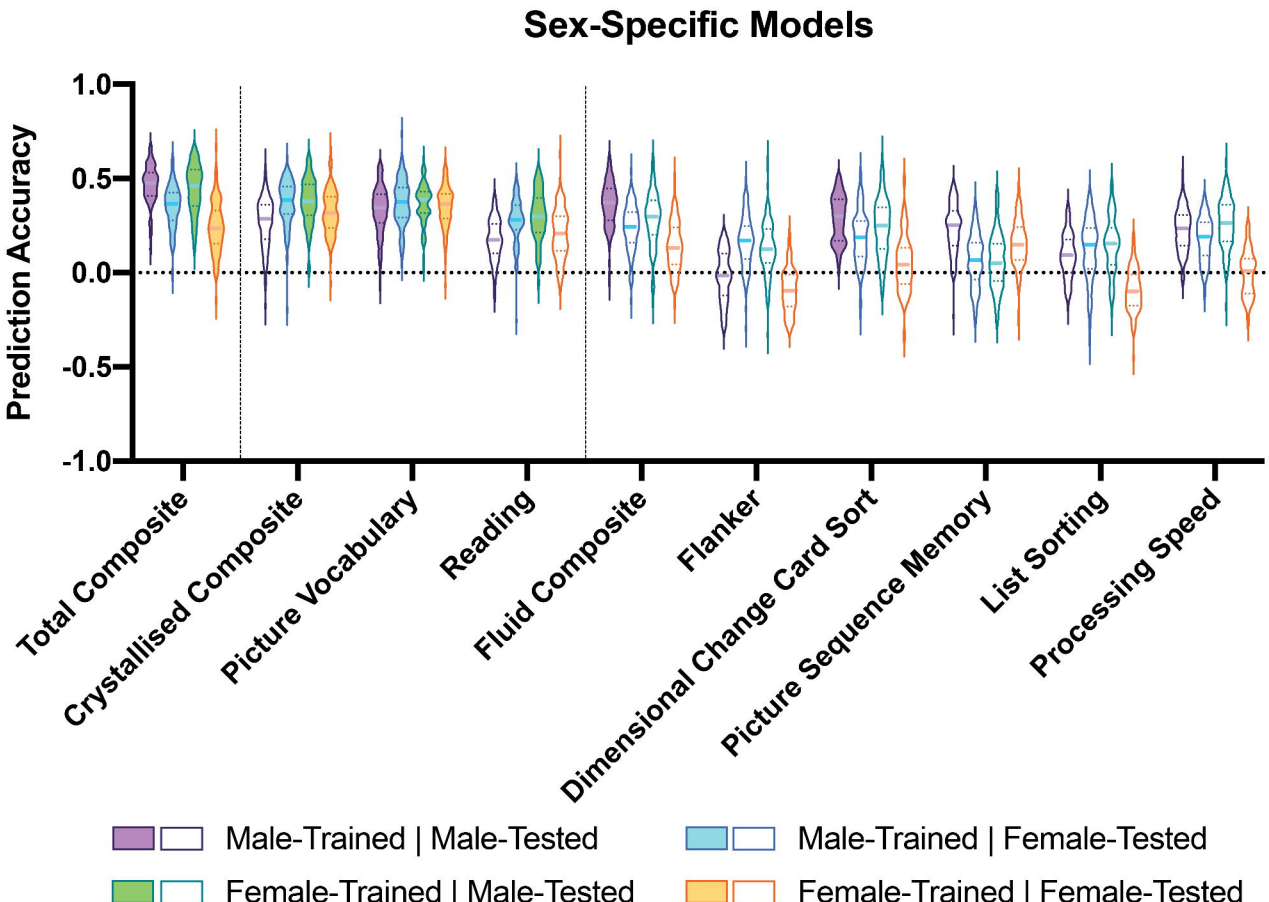
NIH Toolbox Cognition Battery Task	Primary Cognitive Domain	Secondary Cognitive Domain
Dimensional Change Card Sort	Executive Function	Cognitive Flexibility
Flanker Inhibitory Control and Attention	Executive Function	Inhibitory Control and Attention
Picture Sequence Memory	Episodic Memory	
Picture Vocabulary	Language	Vocabulary
Oral Reading Recognition	Language	Reading Decoding
List Sorting Working Memory	Working Memory	
Pattern Comparison Processing Speed	Processing Speed	

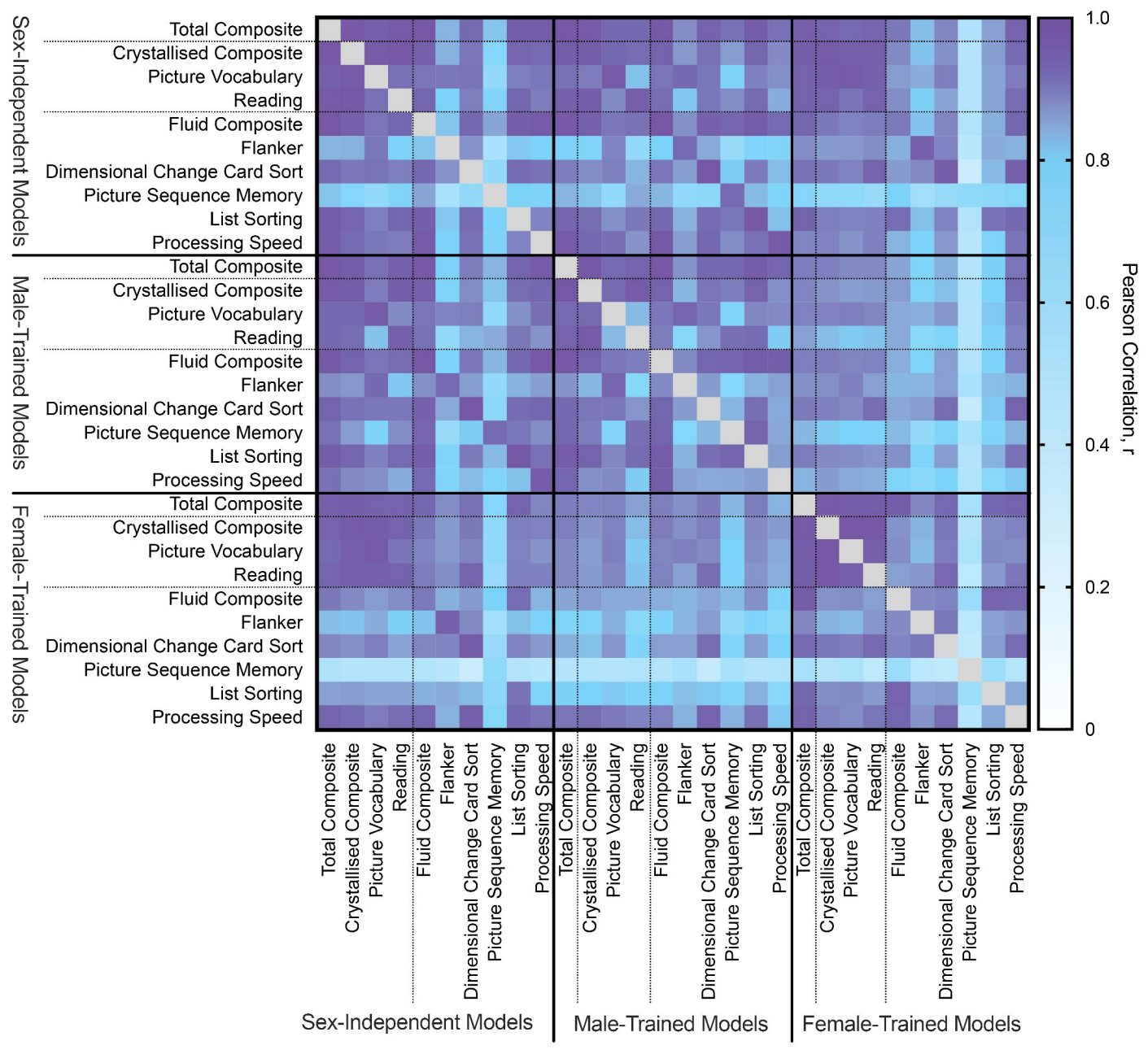
Fluid Cognition	Crystallised Cognition	Total Cognition

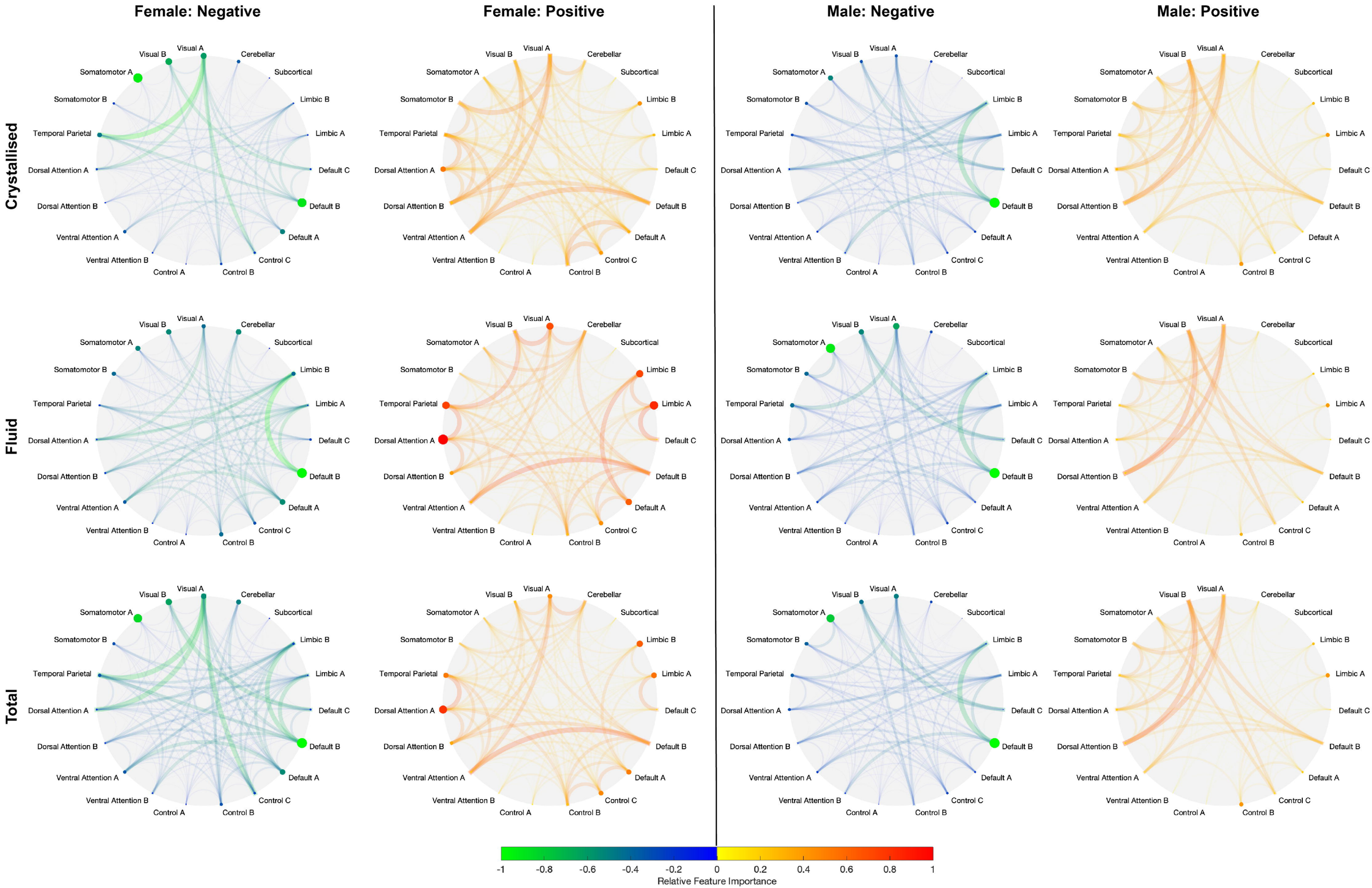
c)

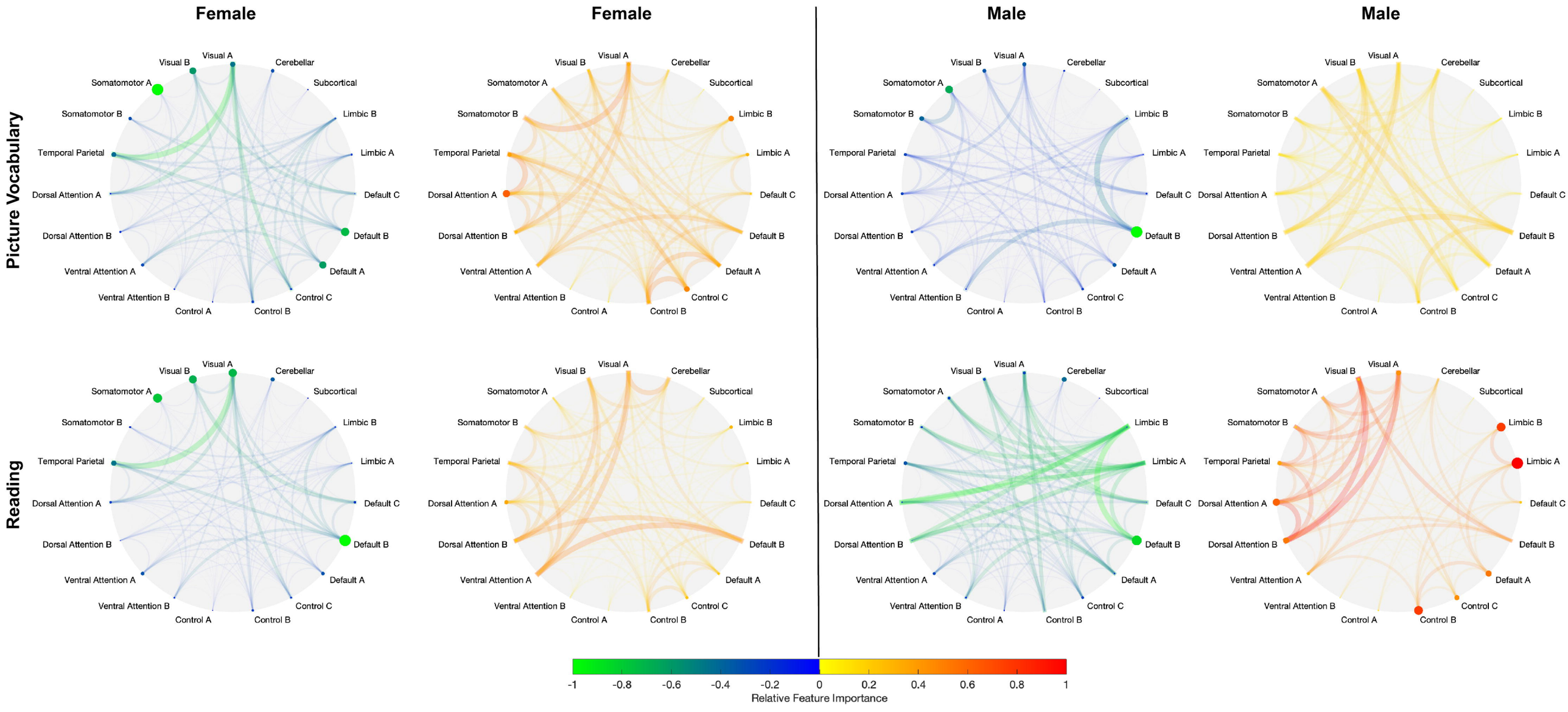



Sex-Independent Models
Prediction Accuracy
1.0
0.5
0.0
-0.5
-1.0
Total Composite
Crystallised Composite
Picture Vocabulary
Reading
Fluid Composite
Flanker
Dimensional Change Card Sort
Picture Sequence Memory
List Sorting
Processing Speed
Male-Tested
Female-Tested


Sex-Specific Models
Prediction Accuracy
1.0
0.5
0.0
-0.5
-1.0
Total Composite
Crystallised Composite
Picture Vocabulary
Reading
Fluid Composite
Flanker
Dimensional Change Card Sort
Picture Sequence Memory
List Sorting
Processing Speed
Male-Trained | Male-Tested
Male-Trained | Female-Tested
Female-Trained | Male-Tested
Female-Trained | Female-Tested


Sex-Independent Models
Male-Trained Models
Female-Trained Models
Total Composite
Crystallised Composite
Picture Vocabulary
Reading
Fluid Composite
Flanker
Dimensional Change Card Sort
Picture Sequence Memory
List Sorting
Processing Speed
Pearson Correlation, r
1.0
0.8
0.6
0.4
0.2
0


Female: Negative
Female: Positive
Male: Negative
Male: Positive
Crystallised
Fluid
Total
Visual A
Visual B
Cerebellar
Subcortical
Limbic B
Limbic A
Default C
Default B
Default A
Control C
Control B
Control A
Ventral Attention B
Ventral Attention A
Dorsal Attention B
Dorsal Attention A
Temporal Parietal
Somatomotor B
Somatomotor A
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
Relative Feature Importance


Female
Female
Male
Male
Picture Vocabulary
Reading
Visual A
Visual B
Cerebellar
Subcortical
Limbic B
Limbic A
Default C
Default B
Default A
Control C
Control B
Control A
Ventral Attention B
Ventral Attention A
Dorsal Attention B
Dorsal Attention A
Temporal Parietal
Somatomotor B
Somatomotor A
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
Relative Feature Importance