

ORIGINAL ARTICLE

Gender Differences in Connectome-based Predictions of Individualized Intelligence Quotient and Sub-domain Scores

Rongtao Jiang^{1,2}, Vince D. Calhoun³, Lingzhong Fan¹, Nianming Zuo¹, Rex Jung⁴, Shile Qi³, Dongdong Lin³, Jin Li¹, Chuanjun Zhuo⁵, Ming Song¹, Zening Fu³, Tianzi Jiang^{1,2,6,7} and Jing Sui^{1,2,3,7}

¹Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China, ²University of Chinese Academy of Sciences, Beijing, 100049, China ³Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, GA 30303, USA, ⁴Department of Neurosurgery, University of New Mexico, Albuquerque, NM 87131, USA, ⁵Department of Psychiatric-Neuroimaging-Genetics and Morbidity Laboratory (PNGC-Lab), Nankai University Affiliated Anqing Hospital, Tianjin Mental Health Center, Tianjin, 300222, China, ⁶University of Electronic Science and Technology of China, Chengdu, 610054, China and ⁷Chinese Academy of Sciences Center for Excellence in Brain Science, Institute of Automation, Beijing, 100190, China

Address correspondence to Jing Sui and Tianzi Jiang, National Laboratory of Pattern Recognition and Brainnetome Center, Institute of Automation, Chinese Academy of Sciences, 95 Zhongguancun East Road, 100190, Beijing, China. Email: kittysj@gmail.com and jiangtz@nlpr.ia.ac.cn.

Abstract

Scores on intelligence tests are strongly predictive of various important life outcomes. However, the gender discrepancy on intelligence quotient (IQ) prediction using brain imaging variables has not been studied. To this aim, we predicted individual IQ scores for males and females separately using whole-brain functional connectivity (FC). Robust predictions of intellectual capabilities were achieved across three independent data sets (680 subjects) and two intelligence measurements (IQ and fluid intelligence) using the same model within each gender. Interestingly, we found that intelligence of males and females were underpinned by different neurobiological correlates, which are consistent with their respective superiority in cognitive domains (visuospatial vs verbal ability). In addition, the identified FC patterns are uniquely predictive on IQ and its sub-domain scores only within the same gender but neither for the opposite gender nor on the IQ-irrelevant measures such as temperament traits. Moreover, females exhibit significantly higher IQ predictability than males in the discovery cohort. This findings facilitate our understanding of the biological basis of intelligence by demonstrating that intelligence is underpinned by a variety of complex neural mechanisms that engage an interacting network of regions—particularly prefrontal–parietal and basal ganglia—whereas the network pattern differs between genders.

Key words: connectome-based prediction, functional connectivity, gender difference, individualized prediction, intelligence

Introduction

Individual differences in general cognitive ability can be measured by the intelligence quotient (IQ), which assesses abilities including planning, reasoning, comprehension, abstraction and learning (Deary 2013). Compared with fluid intelligence (gF), IQ is a more generalizable and complex construct, which is also strongly predictive of various important life outcomes including educational achievement, occupational attainment, social mobility and job performance (Deary et al. 2007).

With the advent of magnetic resonance imaging (MRI), substantial progress has been made in understanding the neurobiological basis of intelligence. The associations between intellectual abilities and neuroimaging measures have been widely reported in gray matter (Ohtani et al. 2014), white matter (Genc et al. 2018), cortical thickness (Narr et al. 2007), and functional connectivity (FC) (Pamplona et al. 2015). Recent work also suggested that, as a complex construct, IQ is underpinned by communications among widespread brain regions (Glascher et al. 2010), including but not limited to the prefrontal and parietal areas, especially the interactions between dorsal attention, fronto-parietal, and default model networks (Hearne et al. 2016; Santarnecchi et al. 2017). A systematic review suggested that from the perspective of Network Neuroscience Theory, the general intelligence originates from the system-wide small-world topology and network dynamic of the human brain, which can enable more flexibility and adaptation in information exchange and processing (Barbey 2018). Currently, there are increasing interests in determining imaging biomarkers that can be used to predict personalized cognitive or health outcomes at the level of an individual (Abi-Dargham and Horga 2016). Particularly, studies have demonstrated that connectome-based predictive models are predictive of various cognitive abilities including sustained attention, personality traits, and divergent thinking (Rosenberg et al. 2016; Yoo et al. 2017; Beaty et al. 2018; Hsu et al. 2018; Jiang, et al. 2018b), owing to the advantage of FC in depicting the interactions between different brain regions (He et al. 2016).

However, only a limited number of studies have attempted to perform quantitative predictions of intelligence scores at the individual level. Specifically, Finn et al. (2015) successfully established the relevance of connectivity profiles to gF by demonstrating its predictability within a fully cross-validated machine-learning analysis, providing a critical foundation for future work to reveal brain-behavior relationships. Besides, a recent study using a cross-validated predictive framework also demonstrated that FCs in a distributed brain network predicted 20% of the variance in general intelligence, which is constructed by factor analysis of the scores on 10 cognitive tasks (Dubois et al. 2018). Moreover, another study applied a similar method to construct a general intelligence factor and showed that the whole-brain task activation maps can serve as a highly effective basis for intelligence prediction (Sripada et al. 2018). Despite such inspiring progress, the field still lacks more evidence demonstrating the predictability of IQ scores using brain imaging measures, and more importantly, clarify on whether the identified intelligence imaging markers can be generalized to new individuals, since most IQ investigations either used limited sample sizes (Arbabshirani et al. 2016) or lack of replication in independent cohorts. By contrast, an effective prediction model for translational neuroscience should be generalizable across contexts and populations, as pointed by a systematical review (Woo et al. 2017). Namely, the trained

models and the identified signatures from discovery cohort are preferably to be able to 1) generalize to independent new individuals; 2) generalize across laboratories, scanners and minor variants in testing conditions; 3) generalize to other outcomes similar to the same target, which can verify specificity of the predictor.

Accordingly, Rosenberg et al. (2016) accomplished robust prediction of individuals' sustained attention task performance using connectome-based predictive modeling, and the identified prediction models successfully generalized to predict a clinical attention metrics and separable components of attention measured by distinct cognitive tasks in completely independent cohorts (Jangraw et al. 2018; Rosenberg et al. 2018). These studies pave the way for developing reliable, robust and generalizable neuroimaging biomarkers of cognitive behaviors that will allow the field to move forward to a translational neuroscience era. The above points inspired us to identify IQ-predictive neuroimaging markers and prediction models that can be generalized across multiple cohorts and contexts with a larger sample size.

On the other hand, sex discrepancy in general intelligence has been a socially and scientifically important topic in cognitive neuroscience because of its prominence in human behavior. Generally, among all cognitive domains of intelligence, males have superior motor and visuospatial abilities, whereas females are supposed to be better at memory and social cognition skills (Halpern et al. 2007; Gur et al. 2012). Moreover, gender differences in the neurobiology of intelligence have also been reported in numerous structural or functional MRI studies (Deary et al. 2010). For example, Schmithorst and Holland (2006, 2007) reported that female subjects have a greater association of intelligence with FC than males, especially connections linking the bilateral Wernicke's areas and left posterior superior temporal gyrus (STG). Narr et al. (2007) indicated that cortical thickness in frontal regions correlates more strongly with intelligence in females, whereas temporal-occipital cortical thickness exhibits a stronger correlation with intelligence in males. Moreover, males' intelligence showed stronger associations with overall white matter volume, whereas females' intelligence demonstrated greater local and global efficiency (Yan et al. 2011; Ryman et al. 2016). In addition, a structural connectome analysis demonstrated that male brains exhibited greater within-hemisphere connectivity and enhanced modularity, while female brains were optimized for interhemispheric communication (Ingalhalikar et al. 2014). These findings motivate our further investigation on gender specificity of the IQ prediction at the individual level.

In the current study, we performed a systematic exploration to determine the predictability of individual IQ scores and the gender difference using whole-brain FCs. Figure 1 demonstrates the analysis flowchart, which encompasses individualized IQ prediction, identification of IQ-predictive FC patterns, as well as evaluation of the gender specificity, intelligence specificity, and generalizability of the identified FC patterns across cohorts. Specifically, first, a fully cross-validated prediction framework incorporating multivariate pattern analysis techniques was applied to a data cohort including 360 college students. Next, gender-specific IQ-predictive FC patterns were estimated for both males and females. Results suggest the IQs of females are more predictable than males when using whole-brain FC, and the identified FC patterns consist of an interacting network including both prefrontal-parietal network and basal ganglia. Furthermore, the identified FC patterns are shown

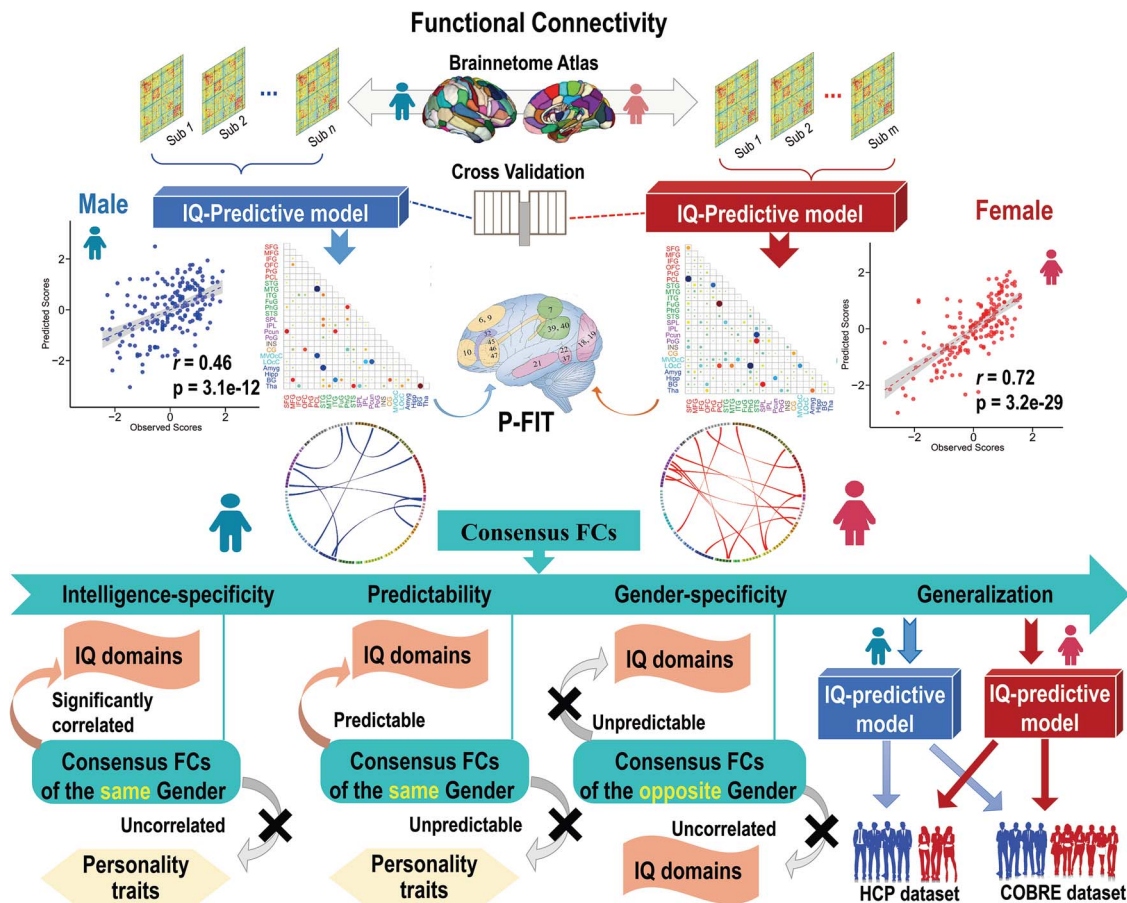


Figure 1. Flowchart of our individualized IQ prediction and validation analysis. In this study, we employed a cross-validated prediction framework to estimate individual's IQ scores using the whole-brain FC. Gender-specific IQ-predictive FC patterns were discovered for both males and females. Moreover, the gender specificity, intelligence specificity, and generalizability of the identified FC patterns were investigated across three independent data sets and two intelligence measurements (IQ and fluid intelligence).

to be uniquely correlated with and predictive on IQ-relevant metrics (e.g. multiple IQ sub-domain scores) but not the IQ-irrelevant measures such as temperament traits. Finally and more importantly, the identified IQ-predictive models can be generalized to two independent cohorts, the Human Connectome Project (HCP) data set with 200 subjects, and the Center of Biomedical Research Excellence (COBRE) cohort with 120 subjects including both healthy controls (HCs) and psychotic disorder over a wide age range. Particularly, all predictions are gender specific, namely, the prediction and generalization work well within the same gender but not the opposite gender.

Materials and Methods

Subjects in Three Data Cohorts

The data from the University of Electronic Science and Technology of China (UESTC) were used as the discovery data set, while data from the HCP Q3 release and the COBRE project were used for validation.

UESTC Data Set

This data set has been used in one of our previous study (Jiang et al. 2018b). A total of 440 healthy college students covering a

wide range of research areas were recruited from the UESTC. All participants were Han Chinese. The research protocol was approved by the Ethics Committee of School of Life Science and Technology at the UESTC. Participants provided written informed consent and were paid for their participation. Participants had no history of neurologic or psychiatric disorders and were not taking any medications that could interfere with their ability to complete a questionnaire or provide MRI data. Each subject was asked to complete the Chinese version of Wechsler Adult Intelligence Scale (WAIS-RC) (Dai et al. 1990). With high test-retest reliability, WAIS is a widely used measurement system that includes several fundamental cognitive performance subtests contributing to intelligence (Jensen 1998) including information, comprehension, digit span, similarities, picture arrangement, block design, and digit symbol. The computed overall score from WAIS, that is Full-Scale IQ, can be used to represent general intellectual abilities. Participants with missing imaging data, incomplete psychological assessment, or excessive head motion (defined as >3 mm translation or $>3^\circ$ rotation during the run) were excluded. Finally, 360 subjects (174 F/186 M, mean age 19.4 ± 1.1 years, in range of 17–24 years) were retained in our study. There is no difference between males and females in age ($P = 0.34$, see Supplementary Fig. S1). In this investigation, Full-Scale IQ scores ranged between 74 and 132 (mean IQ = 108.2 ± 11.4).

HCP Data Set

We used the Q3 HCP data release. A total of 236 healthy subjects were included in this release. After excluding subjects with either missing functional MRI (fMRI) data or missing intelligence cognitive scores, 200 subjects (125 females; 22–36 years, mean age = 29.1 years) were retained. Males and females did not differ in age ($P=0.08$) or education ($P=0.95$). HCP participants did not receive the same intelligence tests as in the UESTC (i.e. WAIS-RC); instead, a different yet related intelligence metrics gF were used. The gF is the capacity to reason and solve novel problems, independent of any knowledge from the past (Jaeggi et al. 2008), which is highly correlated with IQ (Snow et al. 1984). Correlations between gF and several cognition abilities and life outcomes like scholastic achievement, socioeconomic success, and health inequalities have been widely reported already (Colom and Flores-Mendoza 2007). In the HCP protocol, gF was assessed using a form of Raven's progressive matrices (RPM) with 24 items (Bilker et al. 2012) (scores are integers indicating number of corrected items). In this study, gF scores ranged from 4 to 24, with mean score 16.5 ± 4.8 .

COBRE Data Set

A total of 120 participants (42 females; 18–65 years, mean age = 38.0 years) were included for validation analysis, in which 9 were diagnosed with bipolar disorder (BP), 51 with schizophrenia (SZ) or schizoaffective disorder and 60 HCs. Females and males are age ($P=0.52$) and education ($P=0.61$) matched. Informed consent was obtained from all subjects. Detailed inclusion/exclusion criteria can be found in [Supplementary File S1](#). The COBRE subjects received the same intelligence test as in the UESTC (i.e. WAIS). Specifically, IQ scores ranged from 65 to 134, with mean score of 106.9 ± 15.1 .

MRI Data Acquisition

UESTC Data Set

MRI scans were performed on an MR750 3.0 Tesla magnetic resonance scanner (GE Healthcare). Resting-state functional imaging data were acquired using a gradient echo, echo-planar-imaging sequence with the following parameters: time repetition (TR) = 2000 ms, time echo (TE) = 30 ms, field of view (FOV) = 240×240 mm², matrix = 64×64 , flip angle = 90° , voxel size = $3.75 \times 3.75 \times 4.0$ mm³, 36 slices, and 245 volumes. A T1-weighted brain volume MRI sequence was subsequently performed with the following parameters: TR = 8.16 ms, TE = 3.18 ms, flip angle = 7° , FOV = $256 \text{ mm} \times 256 \text{ mm}$, voxel size = $1 \times 1 \times 1$ mm³, and 188 slices. Before scanning, all subjects were instructed to move as little as possible, keep eyes closed, think of nothing in particular, and not fall asleep. Subjects were asked right after the scan whether they had fallen asleep during the scan.

Resting-state fMRI Data Preprocessing

For UESTC data set, imaging data were preprocessed using Data Processing Assistant for Resting-State fMRI Advanced Edition (<http://rfmri.org/DPARSF>). The first 10 volumes were discarded to allow for magnetization equilibrium. Subsequent data preprocessing included slice timing correction, head motion correction, spatial normalization to the Montreal Neurological Institute (MNI) template, resampling to $2 \times 2 \times 2$ mm³, smoothing using a 4 mm Gaussian kernel, temporal band-pass filtering (0.01–

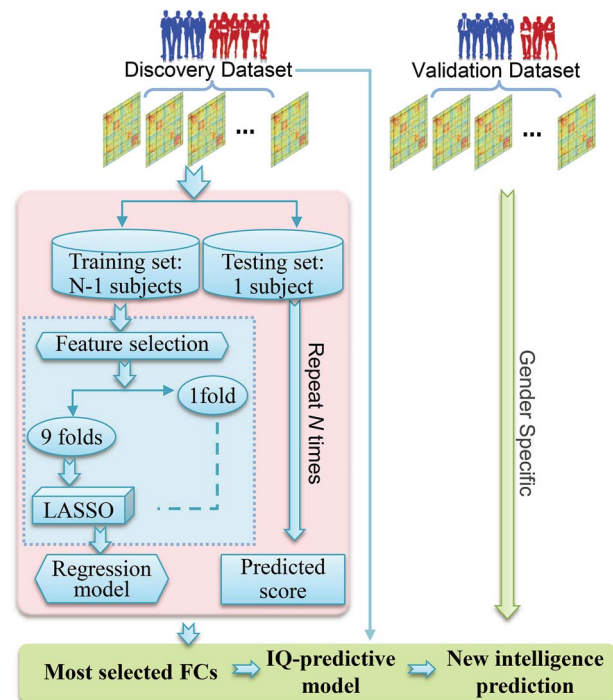


Figure 2. The prediction and validation flowchart incorporating feature selection and regression analysis.

0.08 Hz), and regressing out nuisance signals of head motion parameters, white matter, CSF, and global signals (Jiang et al. 2018b). Details regarding data acquisition and preprocessing for HCP and COBRE cohorts can be found in [Supplementary File S2](#).

Whole-brain FC Extraction

For all three data sets, the preprocessed functional MRI data were parcellated using Brainnetome Atlas (Fan et al. 2016) (https://www.nitrc.org/projects/bn_atlas), resulting in 246 regions of interests (ROIs) that can serve as 210 cortical and 36 subcortical nodes for calculating fine FCs (Jiang et al. 2017). Time series within each node was obtained for each individual by averaging fMRI time series over all voxels in each of the 246 ROIs. Then, Pearson correlations of time courses between each node pair were calculated and were normalized to Z scores using Fisher transformation, resulting in a 246×246 symmetric FC matrix for each subject. After removing 246 diagonal elements, we extracted the upper triangle elements of the FC matrix as the features for prediction, namely, each subject has a FC vector in the dimension of $(246 \times 245)/2 = 30135$.

Individualized Prediction

A prediction framework integrating feature selection and sparse regression was adopted to predict the IQ scores. For neuroimaging data with feature dimension considerably overwhelming the sample size, feature selection is necessary to reduce the redundancy, simplify the fitted model, and enhance generalization (Bermingham et al. 2015). As shown in Figure 2, we employed a nested 10-fold and leave-one-out cross-validation (LOOCV) strategy. Details can be found in [Supplementary File S3](#). Specifically, in each outer loop of the LOOCV, one subject was left

out as testing subject, and the remaining $N-1$ subjects were used as the training set, in which a 10-fold least absolute shrinkage and selection operator (LASSO) regression was run to build the model, where N is the number of subjects (Cui et al. 2016). The details are shown below:

Inner Training Loop

- (i) **Feature selection:** With the ReliefF algorithm (Robnik-Sikonja and Kononenko 1997), every training feature was assigned with a weight statistically accounts for its relevance to the predicted measure, that is the IQ. By determining top m weighted features, we can exclude redundant features effectively. Notably, the determination of m follows existing studies (Dosenbach et al. 2010; Greene et al. 2018; Jiang, et al. 2018a; Liu et al. 2018). [Supplementary File S3](#) provides details on the determination of m and other parameters related to ReliefF.
- (ii) **Regression model building:** The specific IQ scores were estimated with the selected FCs by regression techniques with 10-fold cross-validation in training data set, resulting in a regression model. To compare the prediction performance of different regression algorithms, a total of four popular linear regression models were used here ([Supplementary File S3](#)). Furthermore, to evaluate the effectiveness of feature selection, prediction procedures without ReliefF were also tested for all four regression models, where the whole-brain FCs were used as predictors ([Supplementary File S3](#)).

Outer Predicting Loop

The one left out subject was then input into the regression model derived with inner training data, generating a predicted IQ score. This loop was repeated N times to test through all subjects. Each time, the predicted IQ score for the left-out subject, the identified FCs, and their corresponding weights in the regression model were obtained. By pulling together all testing subjects across N loops, we obtained the predicted IQ scores for all subjects. Pearson's correlations between the predicted and true IQ scores were used to assess predictive power. Moreover, the root mean square error (RMSE) and normalized RMSE (NRMSE) were also calculated (Meng et al. 2017). In order to confirm the specificity of the predictive models and control for potential confounds, the partial correlation between predicted and true IQ scores after ruling out age and mean frame-to-frame displacement was calculated (Cui et al. 2017; Jiang et al. 2018).

Selecting IQ-predictive Consensus FCs

Since we applied a cross-validation strategy to estimate the IQ scores, in each iteration slightly different FCs were selected. For better interpretation and visualization, we grouped the 246 FC nodes into 24 relatively larger brain regions anatomically defined by the Brainnetome atlas (Fan et al. 2016; Rosenberg et al. 2016) and estimated the contributing power of all FCs connecting among these macroscale regions by averaging the regression coefficients of all loops. To circumvent the influence of disproportion of nodes incorporated in different macroscale regions, mean contributing weights were calculated by averaging the total weights of the selected FCs connecting each pair of this macroscale regions. The fundamental features that are common across two gender groups were defined as the shared FCs connecting the same pair of macroscale regions among the

top 100 weighted FCs from males and females, respectively. By selecting the FCs that were repeatedly identified by all loops (i.e. with a 100% occurrence rate), we obtained the "consensus FCs" (Dosenbach et al. 2010; Liu et al. 2018).

Intelligence Specificity of the Consensus FCs

The intelligence-specificity of the identified consensus FCs for each gender was tested by calculating correlations between the consensus FCs and IQ, six intellectual sub-domain, and three temperament traits [harm avoidance (HA), novelty seeking (NS), and reward dependence (RD)] scores (Cloninger et al. 1993), both in the same and the opposite gender group of UESTC. In the traditional view of psychology, there is no meaningful relationship between temperament and intelligence (Zeidner 1995). If the consensus FCs are intelligence specific, then they should show higher correlations with intellectual metrics instead of the temperament traits.

Predictability and Gender Specificity of the Consensus FCs

Note that [Figure 3](#) was unbiased prediction results that include about 100–150 FCs in each of the cross-validations. We further evaluated the predictability and gender specificity of the selected much less "consensus FCs" (accounting for <0.05% of the whole-brain 30 135 FCs) on the above mentioned 10 behavioral metrics. If these consensus FCs can achieve acceptable accuracy on prediction for intelligence-relevant metrics within discovery cohort, then we are able to use them to perform external validations. Specifically, multiple linear regression with rigorous 10-fold cross-validation was performed, in which 8 male-specific or 13 female-specific consensus FCs were used as regressors to predict each of the 10 behavioral metrics in either same gender or opposite gender group. The process was repeated 100 times with subjects randomly shuffled for each metric, and the predictive performance was measured by the mean of the 100 correlations. As hypothesized, if the consensus FCs are gender specific and IQ predictive, then the female-specific FC patterns should be only able to predict intelligence metrics for females instead of males rather than temperament traits in both gender. For males, it is vice versa.

Validation in External Data Sets

We further tested the predictive potential of the IQ-predictive models by conducting external validations in two independent cohorts. In discovery data set, by fitting the IQ scores with the corresponding consensus FCs using multiple linear regression, we can acquire the female-specific and male-specific IQ-predictive models. It is worth noting that predicting with consensus FCs within the discovery data set may cause certain bias due to the fact that features adopted here were derived from the prediction process that incorporated all subjects. The reason for this test here is to ensure that the consensus FCs have the comparable IQ-predictive power and are gender specific, while their reproducibility and generalizability were solely validated in two external data sets as below.

To test the generalizability, we extracted the same FCs as selected in IQ-predictive models from HCP or COBRE subjects through Brainnetome atlas, which were then directly fed into the above IQ-predictive models. Additionally, we calculated the partial correlations between the predicted and observed IQ

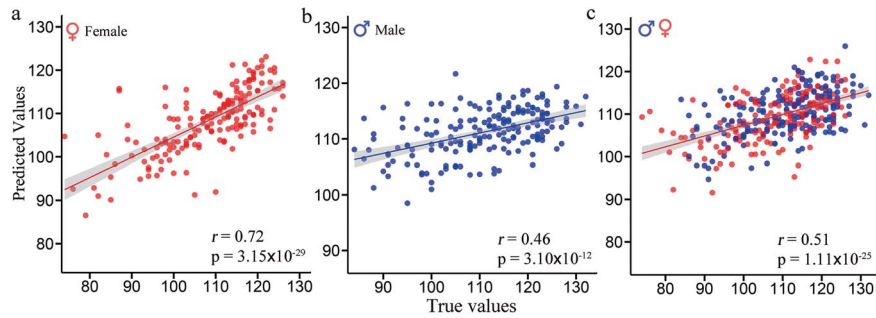


Figure 3. Scatter plot of the predicted IQ scores with respect to their true values for males, females, and all subjects. Based on the prediction framework using whole-brain FC, results revealed significant correlations of $r = 0.72$ ($P = 3.15 \times 10^{-29}$), $r = 0.46$ ($P = 3.10 \times 10^{-12}$), and $r = 0.51$ ($P = 1.11 \times 10^{-25}$) between the predicted IQ scores and true values for females (a), males (b), and all subjects (c).

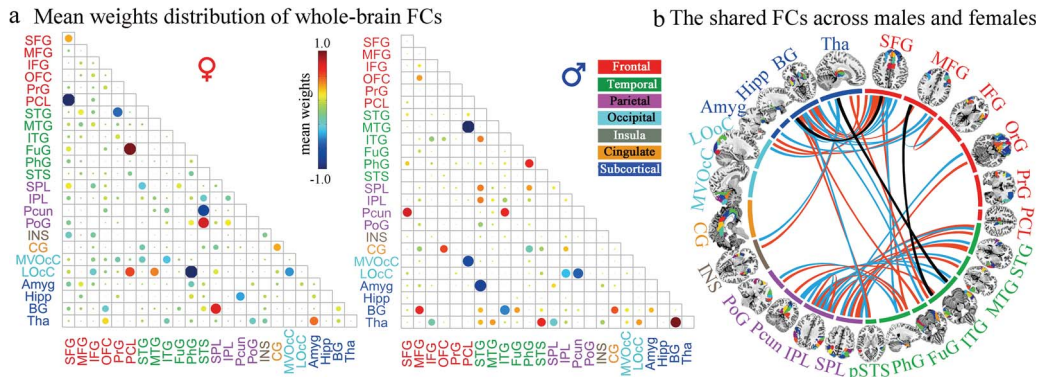


Figure 4. Mean weights distribution of whole-brain FCs and the shared FCs across gender groups. (a) The mean contributing weights of whole-brain FCs for males and females were calculated by averaging the total beta weights in all regression models of the selected FCs. As shown in the matrix plot, the 246 FC nodes are grouped into 24 macroscale brain regions that are anatomically defined by the Brainnetome atlas; Matrix plots: rows and columns represent predefined macroscale regions, and bigger circles represent higher predictive weight. (b) The shared FCs that connect the same pair of macroscale regions among top 100 weighted FCs for females (red line) and males (blue line). Black ones indicate FCs connecting exactly the same pair nodes. As shown in the circle plots, the 246 FC nodes (inner circle) are also grouped into 24 macroscale brain regions (outer brain representations), and nodes incorporated in each of 24 macroscale brain areas are plotted with different colors, which delineate their corresponding anatomy locations in the outer brain representations.

scores by adopting age and mean frame-to-frame motion as control measurements (Hsu et al. 2018).

Results

IQ Prediction in Discovery Data Set

The achieved Pearson correlations between true and predicted IQs were $r = 0.72$ for females (Figure 3a: $P = 3.15 \times 10^{-29}$, $RMSE = 7.9$, $NRMSE = 0.07$), $r = 0.46$ for males (Figure 3b: $P = 3.10 \times 10^{-12}$, $RMSE = 9.7$, $NRMSE = 0.09$), and $r = 0.51$ for all subjects (Figure 3c: $P = 1.11 \times 10^{-25}$, $RMSE = 9.6$, $NRMSE = 0.09$). Notably, as demonstrated in most prediction studies using multivariate regression methods (Rosenberg et al. 2016; Siegel et al. 2016; Greene et al. 2018; Yip et al. 2019), the predicted range was narrower than the observed range; therefore, the model may be most successful at generating predictions of IQ level relative to other subjects (Finn et al. 2015). A Steiger's z-test for testing differences between two independent correlations (Steiger 1980; Fong et al. 2018) revealed that the prediction performance of IQs for females was significantly higher than that for males ($z = 3.86$, $P = 0.0001$). Predictions remain significant even after controlling for age and mean frame-wise head motion (see Supplementary Table S1). Further analysis demonstrated that the difference in predictability was a reflection of actual gender difference, which was not influenced

by data divisions (see Supplementary Fig. S2). Additionally, for robustness, we compared the effectiveness of feature selection, different feature numbers, and four regression models (see Supplementary Table S2 and Fig. S3). Note that female IQ was always more predictable than males regardless of the feature number used and the types of feature selection in all four regression models (see Supplementary Table S2).

IQ-Predictive FC Patterns

Figure 4a demonstrated the mean contributing weights of whole-brain FC in the prediction of IQ scores, with bigger dots representing higher predictive power. Specifically, for females, FC connected regions that show more contributing power predominantly concentrated on superior frontal gyrus [SFG, Brodmann area (BA) 6, 9, 10], precuneus (BA 7), fusiform gyrus (FuG, BA 37), parahippocampal gyrus (BA 35/36), and superior and inferior parietal lobule (BA 7, 39, 40). While for males, FC related regions including middle and inferior temporal gyrus (ITG, BA 20, 21, 37), STG (BA 22, 38), amygdala, basal ganglia, thalamus, and precuneus exhibited more predictive power. Moreover, a further analysis indicated that the direction of these identified feature weights also relates to the true direction of the correlations between FCs and IQ scores (see Supplementary Fig. S4).

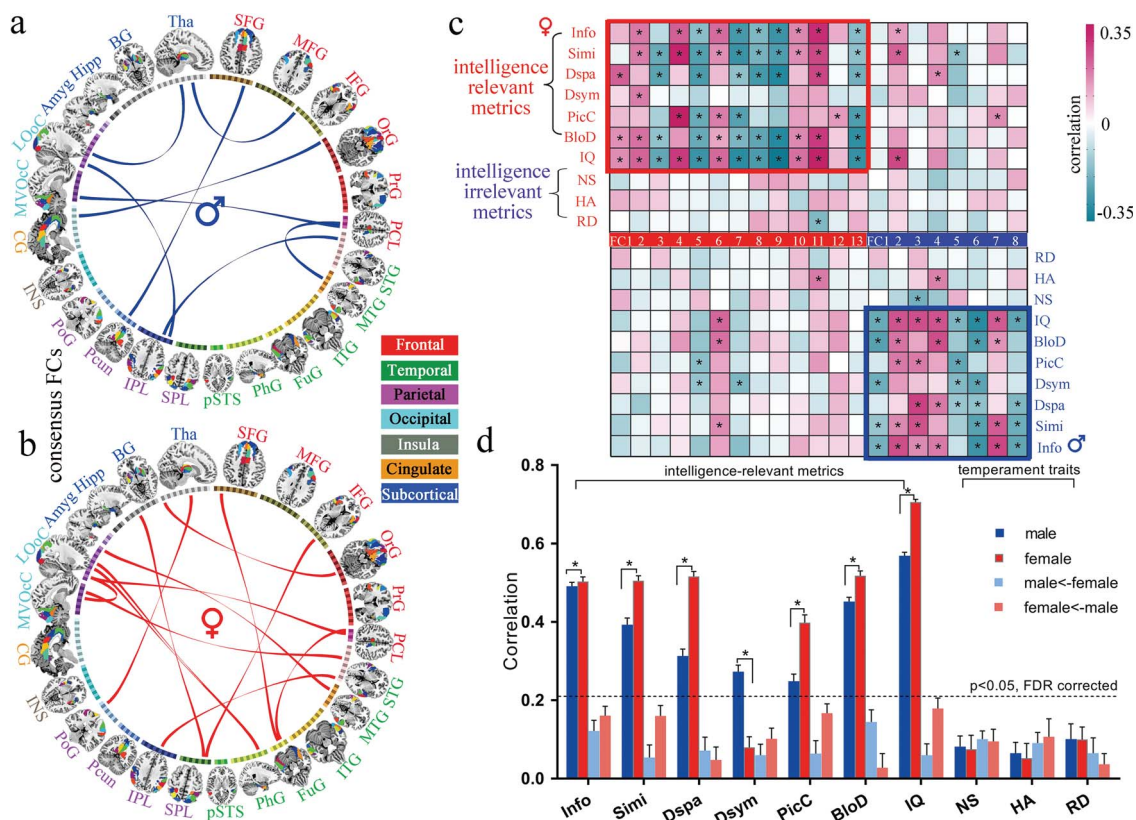


Figure 5. Specificity and predictability of the consensus FC patterns. When summarizing the FC occurrence in all cross-validations, 8 FCs were repeatedly identified for males (a) and 13 for females (b), with a 100% identification rate, and we defined them as the consensus FCs. (c) Correlations between the consensus FCs and IQ, six intellectual sub-domains and three temperament traits scores, both in the same and the opposite gender group (significant correlations with $P < 0.05$ were marked with *). (d) Prediction results for 10 behavior metrics (7 intelligence and 3 temperament) scores were solely based on the consensus FCs from the same or opposite gender group for male and female subjects, respectively. Here we ran multiple linear regression with 10-fold cross-validation, in which 8 male-specific or 13 female-specific consensus FCs were used as regressors to predict each of the 10 behavioral metrics. The process was performed with 100 bootstrapping repetitions with subjects randomly shuffled for each of the 10 behavior metrics. Prediction performance for females with the corresponding consensus FCs are visualized in red and males in blue, while prediction results for females with consensus FCs from the opposite gender are visualized in light red and males in light blue.

Furthermore, to compare between two gender groups, we demonstrated the shared FCs that connect the same pair of macroscale regions among top 100 weighted FCs for females (red line) and males (blue line) among all 30 135 FCs in Figure 4b. Specifically, three gender-common FCs connecting exactly the same nodes pairs (black line) were identified, including FCs between right inferior frontal junction and right caudolateral ITG, right lateroventral ITG, and left rostral temporal thalamus and between left dorsolateral SFG and right ventral caudate. The shared FC-related regions prominently encompassed the prefrontal–parietal regions, ITG, and STG, which were implicated in the parieto-frontal integration theory (P-FIT) inference (Jung and Haier 2007) on human intelligence.

The Identified Consensus FCs

When summarizing the FC occurrence in all cross-validations, 8 FCs were repeatedly identified as consensus FCs for males (Figure 5a) and 13 for females (Figure 5b), with a 100% identification rate, which accounted for less than 0.05% of the whole-brain's 30 135 FCs and are key features for the following validation on IQ predictability and gender specificity. Detailed neuroanatomy information of the consensus FCs and their correlations with IQ scores can be found in Supplementary Table S3.

Intelligence Specificity of the Consensus FCs

As shown in Figure 5c, most of the correlations between each of the consensus FCs and seven intelligence-relevant (IQ and six sub-domains) scores were significant within same gender (red block and blue block) but not significant for three temperament trait scores. More interestingly, few significant correlations were observed between any of the 10 behavioral metrics (7 intelligence and 3 temperaments) and consensus FCs derived from the opposite gender. Particularly, all 8 male-specific consensus FCs were significantly correlated with IQ scores [$P < 0.05$, false discovery rate (FDR) corrected for multiple comparisons], and 12 out of 13 female-specific consensus FCs show significant correlations with IQ scores ($P < 0.05$, FDR corrected).

Gender Specificity of the Consensus FCs

Figure 5d demonstrates the mean prediction accuracies for seven intelligence-relevant metrics and three temperament traits in four combinations of gender prediction. Specifically, significant correlations of $r = 0.57 \pm 0.009$ ($p < 10^{-16}$, RMSE = 9.06 ± 0.07) and $r = 0.71 \pm 0.01$ ($p < 10^{-27}$, RMSE = 8.02 ± 0.09) were revealed for males and females between the predicted and observed IQ scores. Regarding the six intelligence sub-domains, only the digit symbol was not significantly predicted

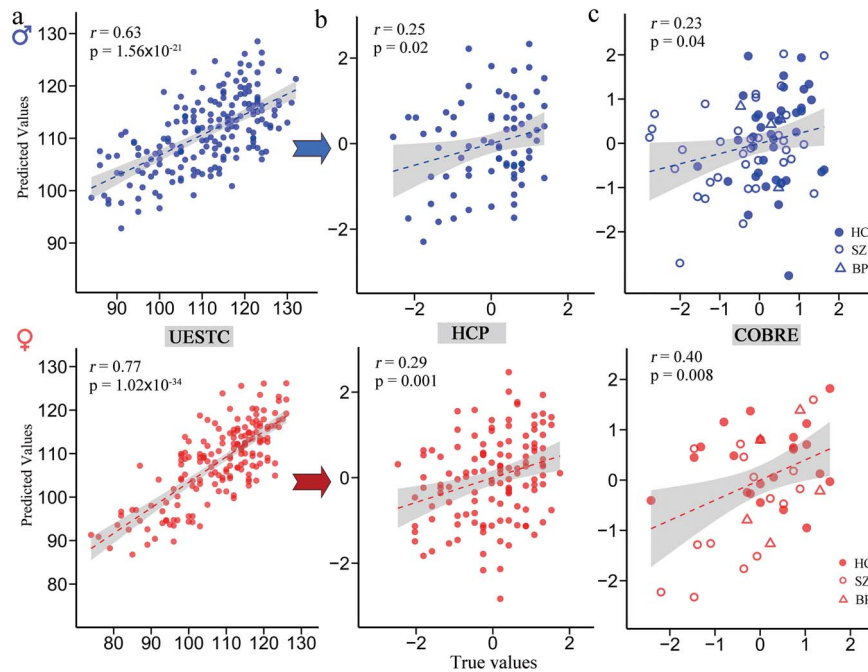


Figure 6. Generalization of the consensus FCs-based predictive models to external data sets. (a) By fitting IQ scores with the corresponding consensus FCs using multiple linear regression in all UESTC male or female subjects, we acquired the male-specific and female-specific IQ-predictive models. The same consensus FCs were extracted from two validation cohorts through Brainnetome atlas and then fed into the models directly to predict (b) the fluid intelligence (gF) for 200 HCP and (c) IQ scores of 120 COBRE subjects. Values in the x-axis and y-axis were normalized for visualization.

by consensus FCs in females ($r = 0.08 \pm 0.03$, $P > 0.1$). As expected, in both gender groups, no significant correlations were achieved for three temperament traits prediction. In addition, there were no significant prediction results in males when predicting any of the 10 behavior metrics with female-specific consensus FCs and neither in females similarly (see [Supplementary Table S4](#)). Moreover, similar to the prediction results using whole-brain FCs, females achieved significantly higher accuracies than males for almost all seven metrics ($P < 0.001$; see [Supplementary Table S4](#)).

IQ Predictability in External Data Sets

By fitting IQ scores with 8 male-specific or 13 female-specific consensus FCs using multiple linear regression for 186 males and 174 females, we acquired the male-specific and female-specific IQ-predictive models in UESTC cohort. Pearson's correlations of $r = 0.63$ ($P = 1.56 \times 10^{-21}$) and $r = 0.77$ ($P = 1.02 \times 10^{-34}$) were achieved for males and females, respectively, which represent the goodness of fit of the gender-specific IQ-predictive models in the UESTC data set ([Fig. 6a](#)). Similarly, female IQ was more predictable than male IQ under Steiger's z-test ($z = 2.62$, $P = 0.009$).

A correlation of $r(73) = 0.253$ ($P = 0.02$) was achieved when applying the predictive model trained in UESTC males to predict the gF scores for 75 HCP males. Similarly, female-specific predictive model achieved a significant correlation of $r(123) = 0.29$ ($P = 0.001$) for 125 HCP female participants ([Fig. 6b](#)). As an even stronger test of generalizability in COBRE data set, which includes subjects with SZ, BP, and HC in a wide range of age (18–65 years), significant correlations were obtained between the predicted and observed IQ scores in 42 females [$r(40) = 0.40$, $P = 0.008$; 22 HC, 5 BP, and 15 SZ] and 78 males

[$r(76) = 0.23$, $P = 0.04$; 38 HC, 4 BP, and 36 SZ; [Fig. 6c](#)]. Predictions remain significant after adjusting for age and mean frame-to-frame head motion, ruling out these potential confounds (see [Supplementary Table S1](#)). However, the difference in prediction performance from males and females did not reach significance in both validation cohorts (HCP: Steiger's $z = 0.29$, $P = 0.39$; COBRE: $z = 0.96$, $P = 0.169$). Interestingly, results were not significant when applying the male-specific or female-specific model to subjects in the opposite gender group for both external validation data sets.

Discussion

An ultimate goal of the use of imaging biomarkers is to perform individualized prediction of educational or health outcomes for new individuals ([Gabrieli et al. 2015](#)), as well as to facilitate our understanding of the neurobiological basis of clinical behaviors. In this work, we carefully investigate the gender difference and generalizability of predicting individualized IQ and sub-domain scores using whole-brain FC. Results indicated a higher predictability for females ([Figure 3, 5d, 6](#), and [Supplementary Table S4](#)) than males, and the gender-specific IQ-predictive models were successfully generalized to two independent data sets, providing evidence for meaningful gender-related heterogeneity in the neurobiology of intelligence.

Summary of Findings

Interestingly, our results initially showed significant gender heterogeneity on IQ predictability, not only on the prediction accuracy ([Fig. 3](#), females: 0.72 vs. males: 0.46) but also on the identified FC patterns ([Fig. 4a](#)). First, based on the prediction framework, we identified several consensus FC patterns (8 for males

and 13 for females; Fig. 5a and b), which show higher correlations with intellectual metrics instead of the temperament traits. Second, using the identified consensus FCs, we successfully predicted the IQ and its six sub-domain scores, achieving high prediction precision for both gender groups, whereas we found no significant results for three intelligence-irrelevant temperament traits. Prediction correlations were comparable with results derived from the whole-brain FCs, suggesting that individual intelligence scores can be predicted solely based on these predictive features. These results suggest that consensus FCs derived from each gender group have great specificity and strong predictability for intellectual abilities. Additionally, predictive models based on these consensus patterns also demonstrated great gender specificity. The female-specific FC patterns were unable to predict any of the intelligence metrics or temperament traits for males well and vice versa. Moreover, specificity of the gender-specific predictive models was also revealed in the validation data set. The female-specific linear model fit to the UESTC data with consensus FCs can be directly applied to predict IQ scores for female subjects in the validation data sets, and the same is true for males. However, we were unable to predict intellectual performance by applying the female- or male-specific models to subjects in the opposite gender group.

Moreover, aiming to identify imaging biomarkers that could guide clinical practice, we tested the IQ-predictive models on their performance to predict outcomes for new individuals and in multiple cohorts. Since the most useful imaging signatures should generalize meaningfully to other outcomes related to the same construct (Woo et al. 2017; Jiang et al. 2018b), in line with this, our current models were developed in a rigorous cross-validation framework, which not only estimated the IQ and its sub-domain scores with a relative high prediction accuracy but also can be generalized to predict multiple intellectual metrics (IQ and gF) successfully across two completely independent data sets (one even include psychotic disorders), suggesting the identified FC patterns are powerful and reliable predictors of intelligence capabilities. Note that IQ (measured by WAIS) and gF [measured by RPM (Bilker et al. 2012)] are two similar but not identical intelligence measurement systems; therefore, the cross-cohort generalization is a powerful evidence to validate the predictability of intellectual performance using the identified connectivity patterns and model (Woo et al. 2017).

Gender Specificity

One most interesting finding is that the identified FC patterns are gender specific and females show more IQ-predictability than males no matter using whole-brain FC features or solely based on consensus FCs. This may be due to different mechanisms underlying the neurobiology of intelligence (Deary et al. 2010). As displayed in Figure 4, for female subjects, the predictive FC nodes prominently concentrated on fundamental P-FIT indicated regions like frontal-parietal networks. However, for males, lingual gyrus and some subcortical areas including amygdala, basal ganglia, and thalamus, beyond those identified in P-FIT (Jung and Haier 2007), were highlighted with more contributing power; while these regions play important roles in a variety of functions including procedural learning, routine behaviors, motivation, reinforcement, decision-making, and working memory. Furthermore, a meta-analysis (Hill et al. 2014) found that although men and women commonly used the same brain networks for working memory, men tended to have a distributed gender-specific networks spread out among the cerebellum,

portions of the superior parietal lobe, and bilateral thalamus. Of note, functional neuroimaging studies have demonstrated a greater correlation between FC and intelligence in females (Schmithorst and Holland 2006, 2007), and such dependence showed an increasing trend in development with age, while the opposite trend was shown in males (Schmithorst 2009). Consequently, the low predictability for males can be partly attributed to a more complex substrate involving more distributed networks in the processing of intelligence leading to a weak correlation between their intelligence and FC.

Moreover, difference in the identified functional connections for males and females may correspond well to their respective superiority in cognitive domains. It has long been recognized that females commonly perform better than males in tasks of verbal abilities and item memory, while males perform better than females in visuospatial processing and mathematical tasks (Bell et al. 2006; Lejbak et al. 2011). For females, the FuG-related functional patterns demonstrated the most predictive power. The FuG, especially the so-called “visual word form area” (VWFA), has been hypothesized to get involved in the lexical processes that bridge the gap between linguistic visual input and speech representations. The VWFA is largely implicated in identifying words and letters from lower-level shape images, prior to association with phonology or semantics, and participating in higher-level processing of word meaning (Dehaene and Cohen 2011). In addition, lesions to the VWFA are associated with impairments in oral reading and oral-naming tasks, as well as reading disorders like the dyslexia (Hillis et al. 2005). For males, FC connected regions consisting of basal ganglia and thalamus were highlighted with more contributing power in the prediction of IQ scores. The caudate nucleus, a main component of the basal ganglia, was supposed to participate in the spatial mnemonic processing (Postle and D’Esposito 2003), which was responsible for the recording of information about one’s environment and spatial orientation. Activity in the caudate nucleus was demonstrated to be greater during tasks featuring spatial and motoric memory demands than those that involved non-spatial tasks (Postle and D’Esposito 1999). Additionally, the basal ganglia (especially the caudate nucleus) has also been reported to play a crucial role in the procedural memory, a core component underlying mathematical abilities (Ullman 2004). Existing studies suggested that abnormalities of brain structures subserving the procedural memory system like basal ganglia could lead to difficulties with math skills (Evans and Ullman 2016). Overall, these results may suggest that males and females commonly take advantage of their most efficient cognitive process in problem solving.

The Parieto-Frontal Integration Theory (P-FIT) and the Identified FC Patterns

Putatively, intelligence is postulated to be underpinned by a distributed network that integrates verbal, visuospatial, attention, working memory, and executive processes, rather than single units (Glascher et al. 2010). One comprehensive synthesis of neuroimaging research on intelligence is the P-FIT, which highlights the significance of an interacting network of brain regions including the prefrontal cortex, parietal cortex, cingulate cortex, and some temporal and occipital regions (Jung and Haier 2007). In our study, multiple brain areas, including basal ganglia, SPL, IPL, STG, ITG, and prefrontal areas were identified as fundamental features with high predictive power, most of which conform to the P-FIT model and correspond to core hubs of

default mode network (DMN), executive control network (ECN), and subcortical network.

The P-FIT has received support from multiple brain lesion studies and analyses of HCs' cognitive data across different neuroimaging modalities (Song et al. 2008; Deary et al. 2010; Glascher et al. 2010; Vakhtin et al. 2014). Moreover, recent work has suggested that, as a complex construct, IQ is underpinned by communication among widespread brain regions, including but not limited to P-FIT areas (Hearne et al. 2016), which conforms with our identified FC patterns. For example, a recent fMRI study reported that high figural creativity, a type of high cognitive functions, was characterized by strong top-down effects between ECN, attention, and memory retrieval networks and weak bottom-up processing between the DMN, subcortical, and primary sensory networks (Kenett et al. 2018; Liu et al. 2018).

Additionally, basal ganglia was revealed in our study as a key FC nodes on IQ prediction, especially caudate nucleus, which is recognized as crucial for learning, particularly when reinforcing or punishing feedback is received contingent on the individual's choice and actions (Tricomi et al. 2006). As known, the ability to learn efficiently is central to most definitions of intelligence. Furthermore, existing studies pointed that the caudate has both afferent and efferent connections to the prefrontal and anterior cingulate cortices, and the dopaminergic tone in these regions is crucial for working memory, which is one of the most important cognitive processes contributing to intelligence (Voorn et al. 2004). Consistent with our findings, Grazioplene et al. (2015) found a positive association between gray matter volume of the bilateral caudate nuclei and intelligence, in three large independent, nonclinical adult samples and linked greater caudate volume to better attentional function, verbal ability, and dopamine receptor availability. Another intelligence study assessed white matter neuroanatomical connectivity found that intelligence showed a significant positive correlation with fractional anisotropy predominantly in the corpus callosum, supporting the idea that efficient information transfer between hemispheres is crucial for higher intellectual capabilities (Navas-Sanchez et al. 2014). Therefore, intelligence is likely to recruit a variety of complex neural mechanisms that engage the whole brain, with emphasis on P-FIT regions (Hearne et al. 2016) but unlike certain cognition processes localized to specific brain regions (Hearne et al. 2016).

Limitations and Future Directions

Some limitations should be considered when interpreting the results. First, despite a high prediction accuracy in discovery data set, the prediction performance in external validation cohorts was relatively low even before being corrected for multiple comparisons, and difference in prediction accuracy between males and females did not reach significance under Steiger's z-test in validation cohorts. This can be probably due to the great heterogeneity of participant populations between discovery and validation cohorts. In our discovery sample, all participants were young, healthy college students with a limited age range (19 to 24 years) for whom higher intellectual capacity than average were likely achieved, given that the developmental effect of age on assessments of intelligence have been widely reported (Cockburn and Smith 1991; Vakhtin et al. 2014). However, participants in HCP (22 to 36 years) and COBRE (18 to 65 years) were much older than those in UESTC and in a relatively lower education level. Apart from HCs, the COBRE cohort also included patients with SZ and BP, which may influ-

ence the sample homogeneity significantly. Moreover, difference in the measurement of intelligence between discovery and validation cohorts can also be a potential reason (IQ vs. gF), given than compared with gF, IQ is a more generalizable and complex construct. Further studies can include more heterogeneous subjects in both discovery and validation cohorts, which could expand the range of observed behavioral performance and add statistical power to the analysis. Second, the cingulate cortex, an ROI implicated in P-FIT findings were absent in our study, which may due to the heterogeneity in our samples who are healthy college students, with a narrower age range compared with other adult neuroimaging studies on intelligence. Finally, the current study was performed using whole-brain FCs from resting-state fMRI, but whether our conclusion holds up using brain connectivity from other cognitive tasks entails being further examined, since a recent study predicting the cognitive behaviors using both resting and task-evoked FC demonstrated that task-induced brain state manipulation improves prediction of individual traits and the task generated the best prediction varies by sex (Greene et al. 2018). For future work, other types of neuroimaging features (structural volume, diffusion weighted imaging, and dynamic FC) can also be employed for prediction of human behavioral measures either separately or in the context of multimodal fusion (Sui et al. 2014; Du et al. 2017; Qi et al. 2018; Sui et al. 2018).

Conclusions

In summary, the current study accomplished robust prediction of IQ scores across multiple cohorts and investigated gender difference for healthy subjects in a relatively large sample using brain FC. The prediction models generalize across three completely independent data sets and two measures of intellectual performance, suggesting that patterns of intrinsic FC could serve as powerful and robust predictors of human intelligence abilities. Interestingly, we found that intelligence of males and females were underpinned by different neurobiological correlates, which are consistent with their respective superiority in cognitive domains (visuospatial vs. verbal ability). Additionally, the identified FC patterns are only uniquely predictive to multiple IQ-relevant metrics scores of the same gender but are not predictive to the opposite gender nor to the temperament traits. This study facilitates our understanding of the biological basis of intelligence by demonstrating that intelligence is underpinned by a variety of complex neural mechanisms that engage numerous brain regions, particularly including those identified by P-FIT and basal ganglia, which however show remarkable gender difference.

Funding

China Natural Science Foundation (No. 61773380); Strategic Priority Research Program of the Chinese Academy of Sciences (grant No. XDB32040100); Brain Science and Brain-inspired Technology Plan of Beijing City (No. Z181100001518005); National Institutes of Health (1R01EB005846, 1R56MH117107, 1R01MH094524, P20GM103472, P30GM122734); National Science Foundation (1539067).

Notes

Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium funded by the 16 National Institutes

of Health (NIH) Institutes and Centers that support the NIH Blueprint for Neuroscience Research and by the McDonnell Center for Systems Neuroscience at Washington University. *Conflict of Interest:* None declared.

Author's Contributions

J.S. and R.J. designed the study; R.J. performed the data analysis; J.S. and R.J. wrote the paper. T.J., V.C., J.L. and L.F. contributed data for analysis. R.J. provided valuable suggestions on IQ knowledge. N.Z., Z.F., D.L., M.S. and S.Q. helped with data preprocessing. All authors contributed to the results interpretation and discussion.

References

- Abi-Dargham A, Horga G. 2016. The search for imaging biomarkers in psychiatric disorders. *Nat Med.* 22:1248–1255.
- Arbabshirani MR, Plis S, Sui J, Calhoun VD. 2016. Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. *Neuroimage.* 145:137–165.
- Barbey AK. 2018. Network neuroscience theory of human intelligence. *Trends Cogn Sci.* 22:8–20.
- Beaty RE, Kenett YN, Christensen AP, Rosenberg MD, Benedek M, Chen Q, Fink A, Qiu J, Kwapił TR, Kane MJ et al. 2018. Robust prediction of individual creative ability from brain functional connectivity. *Proc Natl Acad Sci USA.* 115:1087–1092.
- Bell EC, Willson MC, Wilman AH, Dave S, Silverstone PH. 2006. Males and females differ in brain activation during cognitive tasks. *Neuroimage.* 30:529–538.
- Birmingham ML, Pong-Wong R, Spiliopoulou A, Hayward C, Rudan I, Campbell H, Wright AF, Wilson JF, Agakov F, Navarro P et al. 2015. Application of high-dimensional feature selection: evaluation for genomic prediction in man. *Sci Rep.* 5 Article number: 10312.
- Bilker WB, Hansen JA, Brensinger CM, Richard J, Gur RE, Gur RC. 2012. Development of abbreviated nine-item forms of the Raven's standard progressive matrices test. *Assessment.* 19:354–369.
- Cloninger CR, Svrakic DM, Przybeck TR. 1993. A psychobiological model of temperament and character. *Arch Gen Psychiatry.* 50:975–990.
- Cockburn J, Smith PT. 1991. The relative influence of intelligence and age on everyday memory. *Journal of Gerontology.* 46:P31–P36.
- Colom R, Flores-Mendoza CE. 2007. Intelligence predicts scholastic achievement irrespective of SES factors: evidence from Brazil. *Intelligence.* 35:243–251.
- Cui Z, Su M, Li L, Shu H, Gong G. 2017. Individualized prediction of reading comprehension ability using gray matter volume. *Cereb Cortex.* 28:1–17.
- Cui Z, Xia Z, Su M, Shu H, Gong G. 2016. Disrupted white matter connectivity underlying developmental dyslexia: a machine learning approach. *Hum Brain Mapp.* 37:1443–1458.
- Dai XY, Ryan JJ, Paolo AM, Harrington RG. 1990. Factor-analysis of the mainland Chinese version of the Wechsler Adult Intelligence Scale (WAIS-RC) in a brain-damaged sample. *Int J Neurosci.* 55:107–111.
- Deary IJ. 2013. Intelligence. *Curr Biol.* 23:R673–R676.
- Deary IJ, Penke L, Johnson W. 2010. The neuroscience of human intelligence differences. *Nat Rev Neurosci.* 11:201–211.
- Deary IJ, Strand S, Smith P, Fernandes C. 2007. Intelligence and educational achievement. *Intelligence.* 35:13–21.
- Dehaene S, Cohen L. 2011. The unique role of the visual word form area in reading. *Trends Cogn Sci.* 15:254–262.
- Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, Nelson SM, Wig GS, Vogel AC, Lessov-Schlaggar CN. 2010. Prediction of individual brain maturity using fMRI. *Science.* 329:1358–1361.
- Du Y, Fryer SL, Fu Z, Lin D, Sui J, Chen J, Damaraju E, Mennigen E, Stuart B, Mathalon DH. 2017. Dynamic functional connectivity impairments in early schizophrenia and clinical high-risk for psychosis. *Neuroimage.* 180:632–645.
- Dubois J, Galdi P, Paul LK, Adolphs R. 2018. A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philos Trans R Soc Lond B Biol Sci.* 37 pii: 20170284.
- Evans TM, Ullman MT. 2016. An extension of the procedural deficit hypothesis from developmental language disorders to mathematical disability. *Front Psychol.* 7:1318.
- Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, Yang Z, Chu C, Xie S, Laird AR et al. 2016. The Human Brainnetome Atlas: a new brain atlas based on connectonal architecture. *Cereb Cortex.* 26:3508–3526.
- Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT. 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat Neurosci.* 18:1664–1671.
- Fong AHC, Yoo K, Rosenberg MD, Zhang S, Li CR, Scheinost D, Constable RT, Chun MM. 2018. Dynamic functional connectivity during task performance and rest predicts individual differences in attention across studies. *Neuroimage.* 188:14–25.
- Gabrieli JD, Ghosh SS, Whitfield-Gabrieli S. 2015. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron.* 85:11–26.
- Genc E, Fraenz C, Schluter C, Friedrich P, Hossiep R, Voelkle MC, Ling JM, Gunturkun O, Jung RE. 2018. Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence. *Nat Commun.* 9:1905.
- Glascher J, Rudrauf D, Colom R, Paul LK, Tranel D, Damasio H, Adolphs R. 2010. Distributed neural system for general intelligence revealed by lesion mapping. *Proc Natl Acad Sci U S A.* 107:4705–4709.
- Grazioplene RG, Ryman S, Gray JR, Rustichini A, Jung RE, DY. 2015. Subcortical intelligence: caudate volume predicts IQ in healthy adults. *Hum Brain Mapp.* 36:1407–1416.
- Greene AS, Gao S, Scheinost D, Constable RT. 2018. Task-induced brain state manipulation improves prediction of individual traits. *Nat Commun.* 9:2807.
- Gur RC, Richard J, Calkins ME, Chiavacci R, Hansen JA, Bilker WB, Loughhead J, Connolly JJ, Qiu H, Mentch FD et al. 2012. Age group and sex differences in performance on a computerized neurocognitive battery in children age 8–21. *Neuropsychology.* 26:251–265.
- Halpern DF, Benbow CP, Geary DC, Gur RC, Hyde JS, Gernsbacher MA. 2007. The science of sex differences in science and mathematics. *Psychol Sci Public Interest.* 8:1–51.
- He H, Yu Q, Du Y, Vergara V, Victor TA, Drevets WC, Savitz JB, Jiang T, Sui J, Calhoun VD. 2016. Resting-state functional network connectivity in prefrontal regions differs between unmedicated patients with bipolar and major depressive disorders. *J Affect Disord.* 190:483.
- Hearne LJ, Mattingley JB, Cocchi L. 2016. Functional brain networks related to individual differences in human intelligence at rest. *Sci Rep.* 6 Article number: 32328.

- Hill AC, Laird AR, Robinson JL. 2014. Gender differences in working memory networks: a BrainMap meta-analysis. *Biol Psychol.* 102:18–29.
- Hillis AE, Newhart M, Heidler J, Barker P, Herskovits E, Degaonkar M. 2005. The roles of the “visual word form area” in reading. *Neuroimage.* 24:548–559.
- Hsu WT, Rosenberg MD, Scheinost D, Constable RT, Chun MM. 2018. Resting-state functional connectivity predicts neuroticism and extraversion in novel individuals. *Soc Cogn Affect Neurosci.* 13:224–232.
- Ingalhalikar M, Smith A, Parker D, Satterthwaite TD, Elliott MA, Ruparel K, Hakonarson H, Gur RE, Gur RC, Verma R. 2014. Sex differences in the structural connectome of the human brain. *Proc Natl Acad Sci U S A.* 111:823–828.
- Jaeggi SM, Buschkuhl M, Jonides J, Perrig WJ. 2008. Improving fluid intelligence with training on working memory. *Proc Natl Acad Sci USA.* 105:6829–6833.
- Jangraw DC, Gonzalez-Castillo J, Handwerker DA, Ghane M, Rosenberg MD, Panwar P, Bandettini PA. 2018. A functional connectivity-based neuromarker of sustained attention generalizes to predict recall in a reading task. *Neuroimage.* 166:99–109.
- Jensen AR. 1998. *The g factor: The science of mental ability*, New York (NY): Praeger.
- Jiang R, Abbott CC, Jiang T, Du Y, Espinoza R, Narr KL, Wade B, Yu Q, Song M, Lin D et al. 2018a. SMRI biomarkers predict electroconvulsive treatment outcomes: accuracy with independent data sets. *Neuropsychopharmacology.* 43:1078–1087.
- Jiang R, Calhoun VD, Zuo N, Lin D, Li J, Fan L, Qi S, Sun H, Fu Z, Song M et al. 2018b. Connectome-based individualized prediction of temperament trait scores. *Neuroimage.* 183:366–374.
- Jiang R, Qi S, Du Y, Yan W, Calhoun VD, Jiang T, Sui J. (2017). Predicting individualized intelligence quotient scores using brainnetome-atlas based functional connectivity. IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP), 2017. Tokyo, Japan: IEEE. 1–6.
- Jung RE, Haier RJ. 2007. The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behav Brain Sci.* 30:135–154 discussion 154–187.
- Kenett YN, Medaglia JD, Beaty RE, Chen Q, Betzel RF, Thompson-Schill SL, Qiu J. 2018. Driving the brain towards creativity and intelligence: a network control theory analysis. *Neuropsychologia.* 118:79–90.
- Lejbak L, Crossley M, Vrbancic M. 2011. A male advantage for spatial and object but not verbal working memory using the n-back task. *Brain Cogn.* 76:191–196.
- Liu Z, Zhang J, Xie X, Rolls ET, Sun J, Zhang K, Jiao Z, Chen Q, Zhang J, Qiu J et al. 2018. Neural and genetic determinants of creativity. *Neuroimage.* 174:164–176.
- Meng X, Jiang R, Lin D, Bustillo J, Jones T, Chen J, Yu Q, Du Y, Zhang Y, Jiang T et al. 2017. Predicting individualized clinical measures by a generalized prediction framework and multimodal fusion of MRI data. *Neuroimage.* 145:218–229.
- Narr KL, Woods RP, Thompson PM, Szeszko P, Robinson D, Dimtcheva T, Gurbani M, Toga AW, Bilder RM. 2007. Relationships between IQ and regional cortical gray matter thickness in healthy adults. *Cereb Cortex.* 17:2163–2171.
- Navas-Sanchez FJ, Aleman-Gomez Y, Sanchez-Gonzalez J, Guzman-De-Villoria JA, Franco C, Robles O, Arango C, Desco M. 2014. White matter microstructure correlates of mathematical giftedness and intelligence quotient. *Hum Brain Mapp.* 35:2619–2631.
- Ohtani T, Nestor PG, Bouix S, Saito Y, Hosokawa T, Kubicki M. 2014. Medial frontal white and gray matter contributions to general intelligence. *PLoS One.* 9:e112691.
- Pamplona GS, Santos Neto GS, Rosset SR, Rogers BP, Salmon CE. 2015. Analyzing the association between functional connectivity of the brain and intellectual performance. *Front Hum Neurosci.* 9:61.
- Postle BR, D’Esposito M. 1999. Dissociation of human caudate nucleus activity in spatial and nonspatial working memory: an event-related fMRI study. *Brain Res Cogn Brain Res.* 8:107–115.
- Postle BR, D’Esposito M. 2003. *Spatial working memory activity of the caudate nucleus is sensitive to frame of reference.* *Cognitive, Affective, & Behavioral Neuroscience.* 3:133–144.
- Qi S, Yang X, Zhao L, Calhoun VD, Perrone-Bizzozero N, Liu S, Jiang R, Jiang T, Sui J, Ma X. 2018. MicroRNA132 associated multimodal neuroimaging patterns in unmedicated major depressive disorder. *Brain.* 141:916–926.
- Robnik-Sikonja M, Kononenko I. 1997. An adaptation of Relief for attribute estimation in regression. In: Fisher DH, editor. *ICML*. Burlington (MA): Morgan Kaufmann. pp. 296–304.
- Rosenberg MD, Finn ES, Scheinost D, Papademetris X, Shen X, Constable RT, Chun MM. 2016. A neuromarker of sustained attention from whole-brain functional connectivity. *Nat Neurosci.* 19:165–171.
- Rosenberg MD, Hsu WT, Scheinost D, Todd Constable R, Chun MM. 2018. Connectome-based models predict separable components of attention in novel individuals. *J Cogn Neurosci.* 30:160–173.
- Ryman SG, Yeo RA, Witkiewitz K, Vakhtin AA, van den Heuvel M, de Reus M, Flores RA, Wertz CR, Jung RE. 2016. Frontoparietal gray matter and white matter efficiency differentially predict intelligence in males and females. *Hum Brain Mapp.* 37:4006–4016.
- Santaracchi E, Emmendorfer A, Tadayon S, Rossi S, Rossi A, Pascual-Leone AJ. 2017. Network connectivity correlates of variability in fluid intelligence performance. *Intelligence.* 65:35–47.
- Schmithorst VJ. 2009. Developmental sex differences in the relation of neuroanatomical connectivity to intelligence. *Intelligence.* 37:164–173.
- Schmithorst VJ, Holland SK. 2006. Functional MRI evidence for disparate developmental processes underlying intelligence in boys and girls. *Neuroimage.* 31:1366–1379.
- Schmithorst VJ, Holland SK. 2007. Sex differences in the development of neuroanatomical functional connectivity underlying intelligence found using Bayesian connectivity analysis. *Neuroimage.* 35:406–419.
- Siegel JS, Ramsey LE, Snyder AZ, Metcalf NV, Chacko RV, Weinberger K, Baldassarre A, Hacker CD, Shulman GL, Corbetta M. 2016. Disruptions of network connectivity predict impairment in multiple behavioral domains after stroke. *Proc Natl Acad Sci U S A.* 113:E4367–E4376.
- Snow RE, Kyllonen PC, Marshalek B. 1984. The topography of ability and learning correlations. In: *Advances in the psychology of human intelligence*. Hillsdale (NJ): Erlbaum. Vol 2, p. 103.
- Song M, Zhou Y, Li J, Liu Y, Tian L, Yu C, Jiang T. 2008. Brain spontaneous functional connectivity and intelligence. *Neuroimage.* 41:1168–1176.
- Sripada C, Angstadt M, Rutherford SJB. 2018. Towards a “treadmill test” for cognition: reliable prediction of intelligence from whole-brain task activation patterns. *BioRxiv.* 412056.
- Steiger JH. 1980. Tests for comparing elements of a correlation matrix. *Psychol Bull.* 87:245–251.

- Sui J, Huster R, Yu Q, Segall JM, Calhoun VD. 2014. Function-structure associations of the brain: evidence from multimodal connectivity and covariance studies. *Neuroimage*. 102:11–23.
- Sui J, Qi S, van Erp TGM, Bustillo J, Jiang R, Lin D, Turner JA, Damaraju E, Mayer AR, Cui Y et al. 2018. Multimodal neuro-markers in schizophrenia via cognition-guided MRI fusion. *Nat Commun*. 9:3028.
- Tricomi E, Delgado MR, McCandliss BD, McClelland JL, Fiez JA. 2006. Performance feedback drives caudate activation in a phonological learning task. *J Cogn Neurosci*. 18:1029–1043.
- Ullman MT. 2004. Contributions of memory circuits to language: the declarative/procedural model. *Cognition*. 92:231–270.
- Vakhtin AA, Ryman SG, Flores RA, Jung RE. 2014. Functional brain networks contributing to the Parieto-Frontal Integration Theory of Intelligence. *Neuroimage*. 103:349–354.
- Voorn P, Vanderschuren LJ, Groenewegen HJ, Robbins TW, Pennartz CM. 2004. Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci*. 27:468–474.
- Woo CW, Chang LJ, Lindquist MA, Wager TD. 2017. Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci*. 20:365–377.
- Yan C, Gong G, Wang J, Wang D, Liu D, Zhu C, Chen ZJ, Evans A, Zang Y, He Y. 2011. Sex- and brain size-related small-world structural cortical networks in young adults: a DTI tractography study. *Cereb Cortex*. 21:449–458.
- Yip SW, Scheinost D, Potenza MN, Carroll KM. 2019. Connectome-based prediction of cocaine abstinence. *Am J Psychiatry*. 176:156–164.
- Yoo K, Rosenberg MD, Hsu WT, Zhang S, Li CR, Scheinost D, Constable RT, Chun MM. 2017. Connectome-based predictive modeling of attention: comparing different functional connectivity features and prediction methods across datasets. *Neuroimage*. 167:11–22.
- Zeidner M. 1995. Personality trait correlates of intelligence. In: Saklofske DH, Zeidner M, editors. *International handbook of personality and intelligence*. Boston: Springer US, pp. 299–319.