A Connectivity-based Psychometric Prediction Framework for Brain-behavior Relationship Studies

Jianxiao Wu[1,2], Simon B. Eickhoff[1,2], Felix Hoffstaedter[1,2], Kaustubh R. Patil[1,2], Holger Schwender[3] and Sarah Genon[1,2]

[1]Institute for Systems Neuroscience, Medical Faculty, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany,
[2]Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Center Jülich, Jülich, Germany,
[3]Mathematical Institute, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany

Correspondence and requests for materials should be addressed to Jianxiao Wu (j.wu@fz-juelich.de) or Sarah Genon (s.genon@fz-juelich.de).

Abstract

The recent availability of population-based studies with standard neuroimaging measurements and extensive psychometric characterization opens promising perspectives to investigate the relationships between interindividual variability in brain regions' connectivity and behavioral phenotypes. However, the multivariate nature of the prediction model based on connectivity within a network of brain regions severely limits the interpretation of the brain-behavior patterns from a cognitive neuroscience perspective. To address this issue, we here propose a connectivity-based psychometric prediction (CBPP) framework based on individual region's connectivity profile. Preliminary to the development of this region-wise machine learning approach, we performed an extensive assessment of the general CBPP framework based on whole-brain connectivity information. Because a systematic evaluation of different parameters was lacking from previous literature, we evaluated several approaches pertaining to the different steps of a CBPP study. We hence tested 72 different approach combinations in a cohort of over 900 healthy adults across 98 psychometric variables. Overall, our extensive evaluation combined to an innovative region-wise machine learning approach, offering a framework that optimizes both prediction performance and neurobiological validity (and hence interpretability) to study brain-behavior relationships.

The relationships between brain regions and behavioral functions is a fundamental question for neuroscience. These relationships can be investigated by relating interindividual variability in brain regions' connectivity to interindividual differences in behavioral performance (that is, to psychometric data). The recent availability of population-based neuroimaging datasets with extensive psychometric characterization (Van Essen et al. 2013; Caspers et al. 2014) opens promising perspectives to investigate the relationships between brain regions' connectivity and behavior. In particular, several studies have shown that individual profile of functional connectivity (FC) between brain regions can predict individual scores on psychometric variables, including cognitive measures such as fluid intelligence, as well as personality traits, such as openness (Fin et al. 2015; Rosenberg et al. 2016; Smith et al. 2016; Noble et al. 2017; Dubois et al. 2018; Li et al. 2019). Given the potential of these approaches for cognitive and clinical neuroscience, developing a scientifically valid and useful framework is a crucial objective for the neuroimaging community.

In a cognitive and clinical neuroscience framework, not only the prediction performance matters, but also the neurobiological validity of the model and relatedly, its interpretability, raising one main issue of the current state of the art. Frequently, researchers try to interpret the model with some type of post-hoc evaluation, looking at the brain connectivity features (i.e. region-to-region connectivity values) seemingly playing an important role in the prediction. The relative relevance of the features is often derived from the weights assigned by the regression algorithms. Such interpretations can be hard when the weights are not sparse. Even when sparsity is enforced, interpreting the weights is misleading as the relative magnitude of these weights does not reflect the magnitude of the regions' association with the given psychometric variable. To address this issue, we here propose a connectivity-based psychometric prediction (CBPP) framework based on individual region's connectivity profile. Such an approach is concretely performed by evaluating a machine learning model predicting psychometric data from brain FC independently for each brain region. The prediction performance of psychometric data of each independent brain region's model can then be used as an indicator of the relationship between the specific brain region's connectivity profile and the measured behavioral function.

As a research field, neuroimaging analyses methods, in particular for probing FC, based on resting-state functional magnetic resonance imaging (fMRI) data is a very quickly growing field. Consequently, most previous studies using FC to predict psychometric outcomes and focusing on a specific neuroscientific question cannot afford an extensive methods assessment. To address this issue, preliminary to the development of our region-wise machine learning approach, we conducted a systematic assessment of the CBPP framework, using resting-state fMRI and psychometric data from the Human Connectome Project (HCP; Van Essen et al. 2013). The general workflow of the evaluated CBPP framework follows standard protocols (Shen et al. 2017) and is illustrated in Figure S1. We here investigated the influence of various FC features, types of processing, the interaction between them, as well as the prediction algorithm on the prediction performance. Concretely, at each major step of the CBPP framework, we examined the impact of common parameters: 1) preprocessing/denoising and data compression (i.e. parcellation granularity); 2) FC estimation; 3) regression methods and confounds removal. In our evaluation, we mainly considered previously used approaches,

summarized in Table 1. For this extensive evaluation, we used the whole brain connectivity matrix as features. Accordingly, we refer to this approach as whole-brain CBPP since all region-to-region connectivity values are initially used in the implementation. Capitalizing on this preliminary investigation, we then examined the performance of our region-based approach, here called parcel-wise CBPP approach. To promote the use of this framework in cognitive neuroscience studies, we illustrated two main applications for which we evaluated the brain region-behavior relationships validity in a neuroscience perspective: 1) single brain region's predictive power for different psychometric variables 2) variation of predictive power across brain regions for a single psychometric variable.

Results

Whole-brain CBPP. To evaluate the impact of common parameters in CBPP, we implemented whole-brain CBPP in which a matrix summarizing all parcel-to-parcel (i.e. region-to-region) connectivity values provides the input features for each individual subject. We then assessed the effect of a combinations of different approaches in CBPP using this whole-brain connectivity information. Figure 1 shows the different approaches considered at each step in our implementation of whole-brain CBPP. In total, 72 combinations of approaches were evaluated with 'standard' confound controlling approach.

We used readily available preprocessed resting-state fMRI time series (N=928 for minimally processed, N=923 for FIX and FIX+GSR) and 98 psychometric variables from the HCP initiative. The Schaefer atlas (Schaefer et al. 2018) was used to group the time series into different number of parcels (i.e. at different parcellation granularity). FC is computed as either the Pearson or partial correlation between the mean time series of every pair of parcels, resulting in, for example, a 400x400 FC matrix at 400-parcel granularity. A set of 9 confounding variables (sex, age, age^2, sex*age, sex*age^2, handedness, brain size, intracranial volume, and acquisition quarter) were regressed out from the psychometric variables for the 'standard' confound controlling approach. Finally, we focused on a selection of linear regression techniques popular in the neuroscience field: multiple linear regression (MLR; Matlab's regress function; Chatterjee and Hadi 1986), linear Support Vector Regression (SVR; Boser et al. 1992; Cortes and Vapnik 1995) and Elastic net (EN; Zou and Hastie 2005). For each combination of approaches, we performed 10 repeats of 10-fold cross-validation. Default parameter was used for linear SVR, while hyperparameter tuning for EN was done by using one training fold as inner validation fold (see Methods).

The final prediction accuracy was measured as the average Pearson correlation between the predicted and observed psychometric values, across all test set folds and all repeats. The 98 psychometric variables were ranked according to the average prediction accuracy achieved across the 72 combinations of approaches (Figure S2). Figure 2 shows the whole-brain CBPP results from all 72 different combinations of approaches. Each point on the line plot shows the average prediction accuracy across the 20 most well predicted psychometric variables (Figure S2), for a specific combination of approaches. Similarly, figure S3 shows the whole-brain CBPP results from all 72 combinations, with the average prediction accuracy across all the 98 predicted psychometric variables.

Numerically, the highest average prediction accuracy was achieved by the SVR-minimal-partial combination at 300-parcel granularity, followed by SVR-FIX-partial combination at 200-parcel and 300-parcel granularities. For combinations using Pearson correlation, the highest average prediction accuracies were achieved by FIX+GSR combinations using SVR and EN, at 200-parcel and 300-parcel granularities.

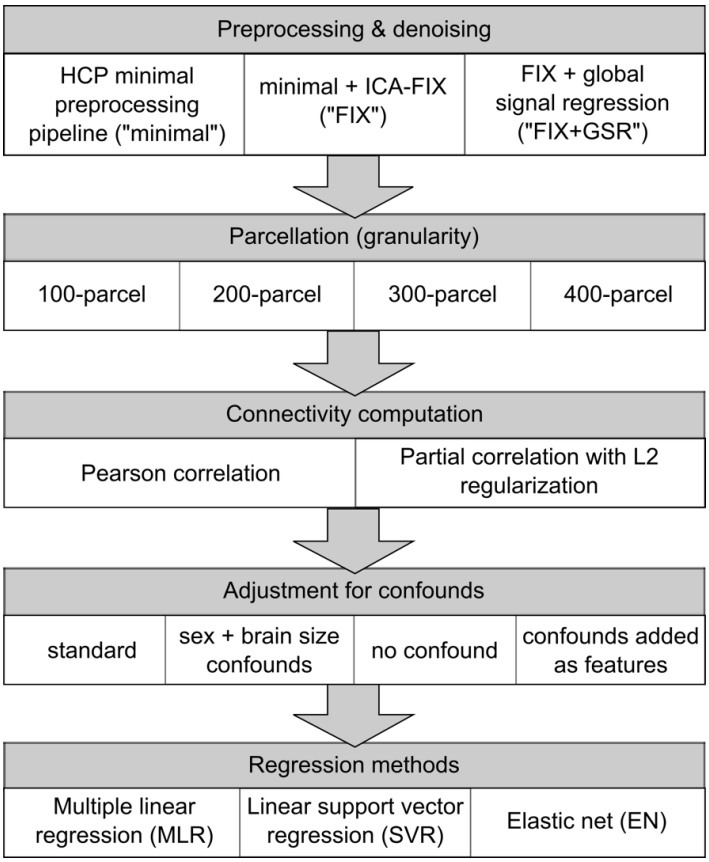

Figure 1. Approaches considered for each step in the CBPP framework.

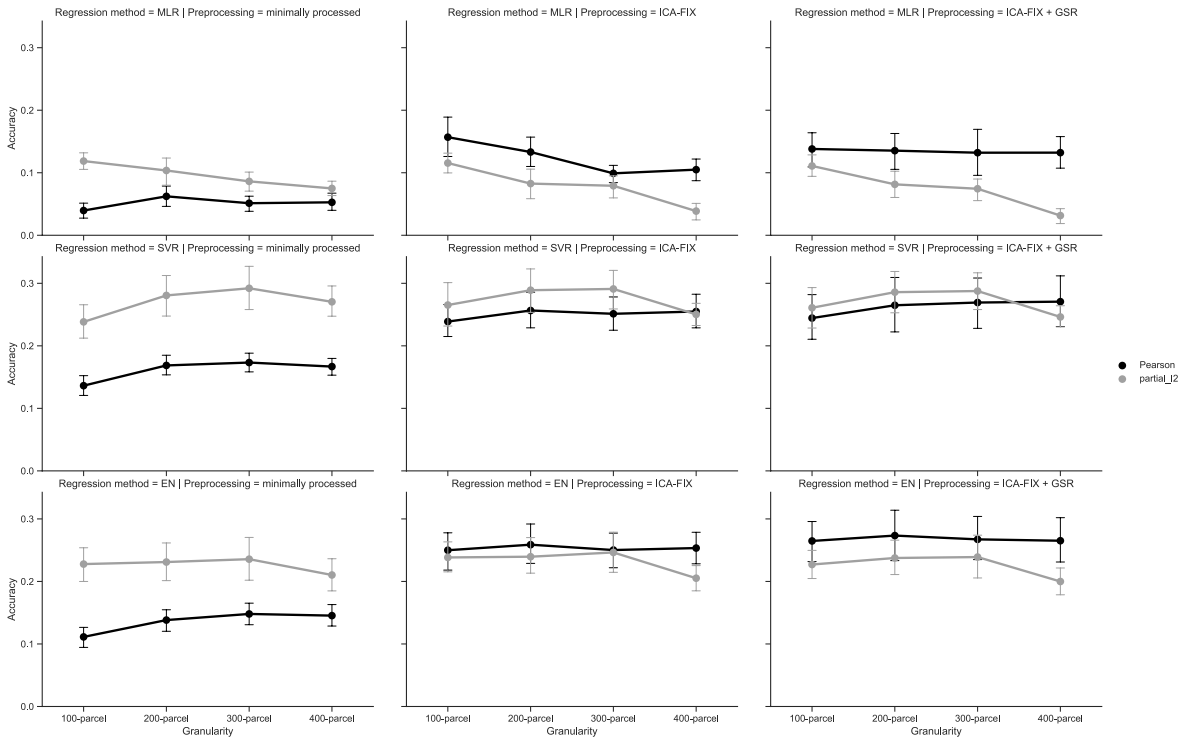

Figure 2. Average prediction accuracy across the 20 most well predicted psychometric variables for each combination of approaches in whole-brain CBPP. Error bars represent the 95% confidence interval across psychometric variables. The columns show combinations using minimally processed, FIX, and FIX+GSR data respectively; the rows show combinations using MLR, SVR and EN respectively. Dark lines represent combinations using Pearson correlation, while light grey lines represent combinations using partial correlation.

Impact of common parameters on whole-brain CBPP. Different trends of prediction accuracy by parcellation granularity were observed when different connectivity computation methods were used. For combinations using Pearson correlation, prediction accuracies were generally observed to increase with granularity. Combinations using partial correlation showed a slightly different pattern in which prediction accuracies increased with granularity until 300-parcel level generally, before dropping at 400-parcel level.

The impact of connectivity computation methods was dependent on the preprocessing methods and regression methods used. For minimally processed data, combinations using partial correlation significantly outperformed those using Pearson correlation ($p < 0.002$ corrected; except MLR-minimal combinations at 300-parcel and 400-parcel granularity). For other preprocessing methods, combinations using SVR always reached quantitatively (but not statistically significantly) higher performance with partial correlation, while combinations using MLR and EN usually reached significantly higher performance with Pearson correlation ($p < 0.03$ corrected).

Preprocessing methods comparisons showed that minimally processed data always led to worse results when Pearson correlation was used ($p < 0.02$ corrected), but were mostly comparable to FIX and FIX+GSR data when partial correlation was used. The comparisons between FIX and FIX+GSR were mostly not statistically significant.

Finally, comparing the regression methods, both SVR and EN outperformed MLR in all combinations ($p < 0.001$ corrected). SVR and EN showed mostly comparable performance when Pearson correlation was used, while SVR outperformed EN when partial correlation was used ($p < 0.01$ corrected; except minimal-partial combination at 100-parcel granularity).

To summarize, machine learning-based denoising (FIX) and a parcellation granularity of 200- or 300-parcel led to the most optimal combinations. Following FIX denoising and 200-parcel/300-parcel granularity, similar prediction performance could be achieved using any connectivity computation method or regression method (except MLR which led to lower performance).

Parcel-wise psychometric profiles. To directly investigate the neurobiological validity and hence interpretability of the brain-behavior patterns, we then examined parcels' psychometric profile with parcel-wise CBPP. The procedure of parcel-wise psychometric prediction was overall the same as that shown in Figure S1 and Figure 1. We implemented parcel-wise CBPP using both minimally processed and FIX data, while focusing on the combination of approaches that showed the best quantitative performance for whole-brain CBPP: minimal-Pearson-SVR and FIX-Pearson-SVR at 300-parcel granularity with the 'standard' confound controlling approach.

Figure 3 and 4 show the psychometric profiles of 4 pairs of parcels, corresponding to parcels in the brain regions primary visual cortex, premotor cortex, supramarginal gyrus and Broca region (according to the Schaefer atlas' labels). Overall, the selected parcels showed similar psychometric profiles across hemispheres, but distinct profiles across parcel locations. Psychometric variables describing general cognitive abilities tended to be well predicted across combinations of approaches in whole-brain CBPP and across parcels in parcel-wise CBPP. For example, the total cognition composite score was the fourth best predicted across combinations of approaches on average and among the best predicted variables for all the 4 selected parcels.

While most of the well predicted psychometric variables for parcels in the primary visual cortex were also well predicted in many other brain regions, processing speed (*Proc Speed*) was predicted with high accuracy only in the primary visual cortex parcels. In contrast, FC patterns of the primary visual cortex parcels showed nearly no predictive power for psychometric variables in the emotion domain.

Both left and right parcels in the premotor cortex showed higher prediction accuracies for working memory performance (*2-back Acc* and *2-back Acc Face*), total cognition composite score (*Total Comp*), cognitive flexibility (*Card Sort*) and endurance (*Endurance*). However, in comparison to other parcels, the premotor cortex parcels showed overall low prediction accuracies for most psychometric variables.

The parcels in the supramarginal gyrus showed high predictive power in most cognition-related variables, including total cognition composite score, fluid intelligence (*Fluid Int*), working memory performance, picture vocabulary (*Pic Vocab*) and crystallized cognition composite score (*Crystal Comp*).

Finally, the parcels in the Broca region showed relatively higher prediction power for language-related measures, cognition composite scores and working memory performance. Interestingly, the right parcel showed lower prediction power for language task accuracy (*Lang task*) than the left, but higher prediction power for working memory abilities.

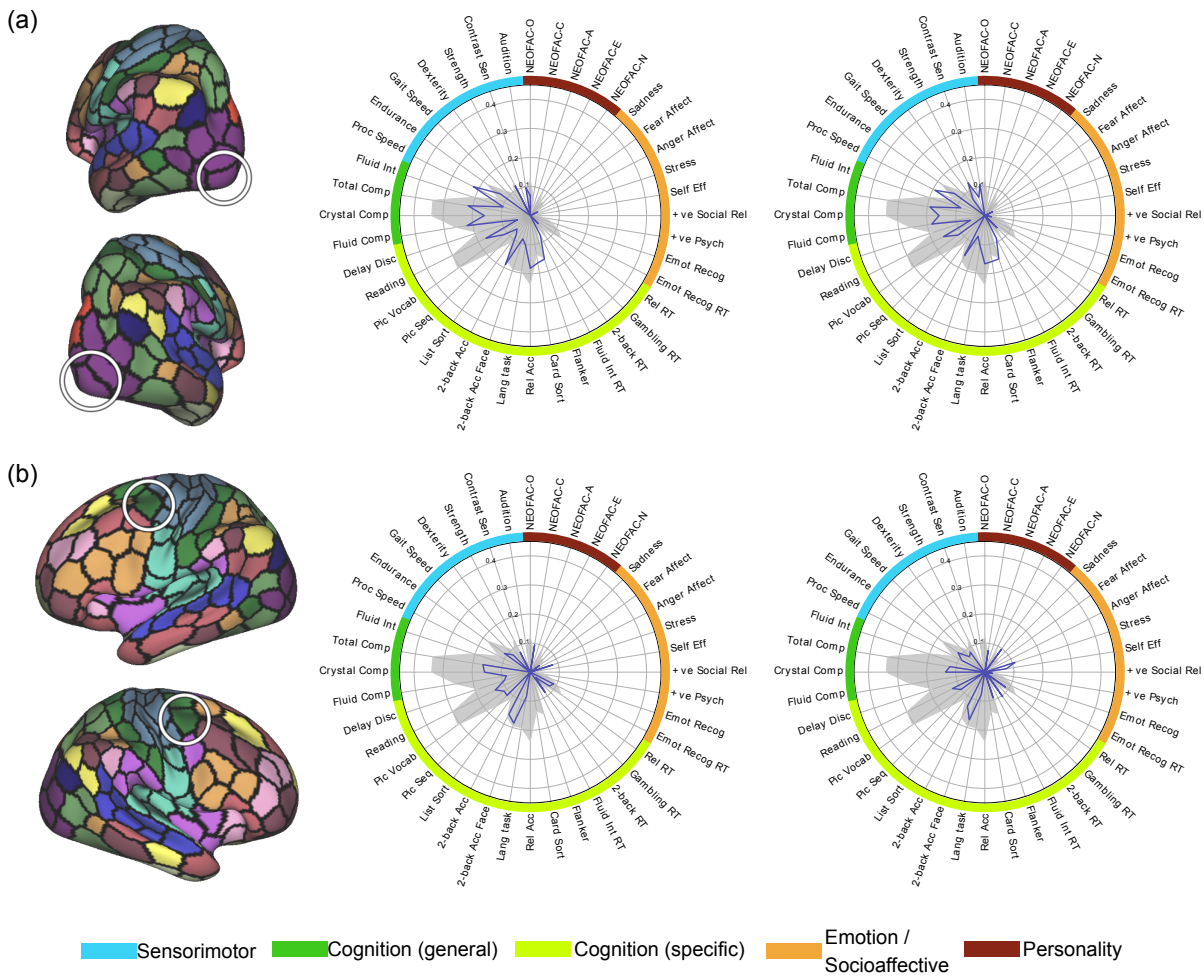

Figure 3. Psychometric profiles for pairs of parcels in (a) Primary visual cortex (b) Premotor cortex in left and right hemispheres respectively, using FIX-Pearson-SVR combination at 300-parcel granularity. Gray filled contour shows whole-brain prediction profile, while blue contour shows parcel-wise prediction profile.

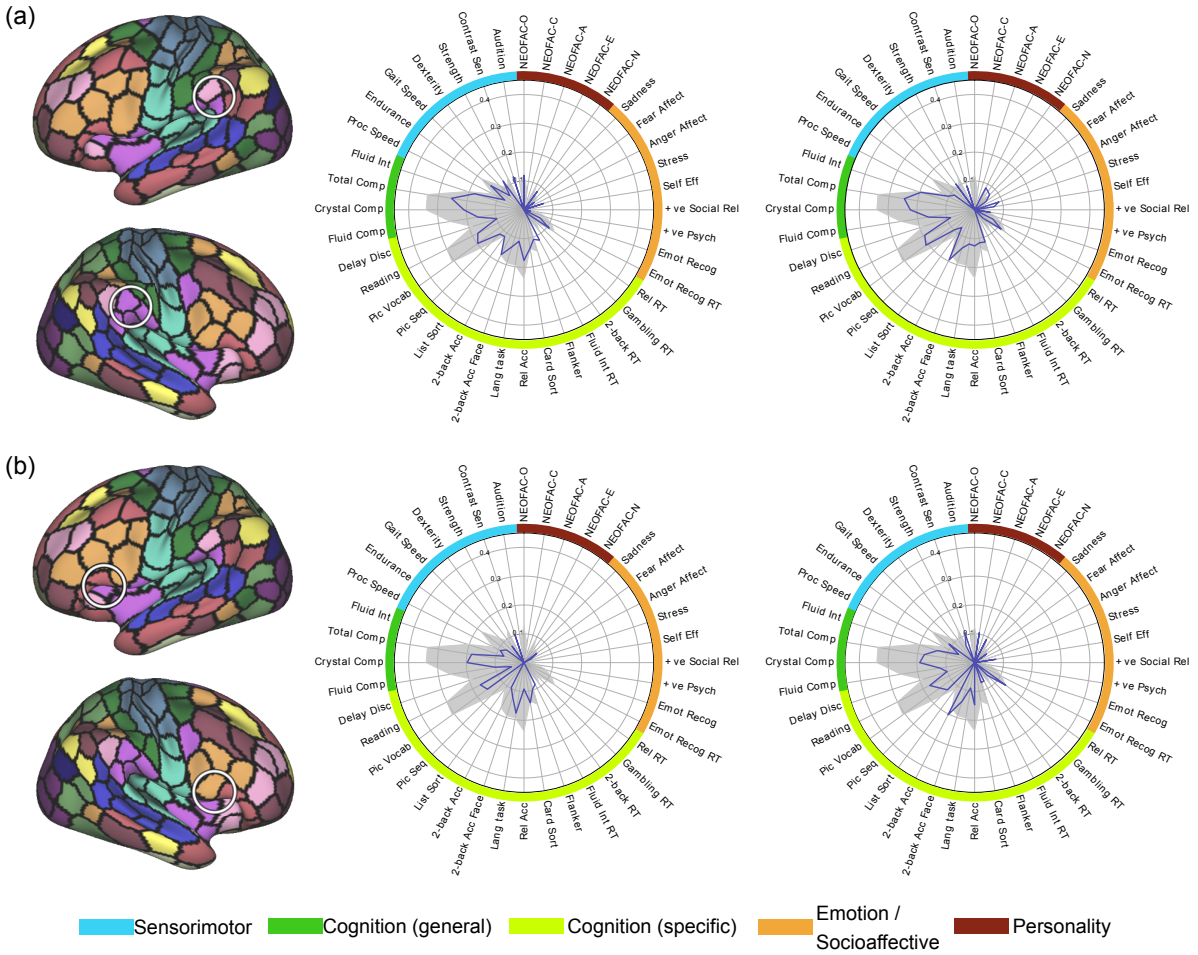

Figure 4. Psychometric profiles for the pair of parcels in (a) Supramarginal gyrus (b) Broca region in left and right hemispheres respectively, using FIX-Pearson-SVR combination at 300-parcel granularity. Gray filled contour shows whole-brain prediction profile, while blue contour shows parcel-wise prediction profile.

Parcel-wise prediction accuracy distribution. From the psychometric variables' perspective, we present the prediction accuracies distribution across parcels. Figure 5a to 5d show the prediction accuracy distributions and histograms across the brain for 4 selected psychometric variables, using FIX-Pearson-SVR combination at 300-parcel granularity. Across the brain, prediction accuracies were generally lower for strength and higher for crystallized cognition composite score.

For strength prediction, the best performing parcel was the parcel in the ventral postcentral sulcus in the left hemisphere. For crystallized cognition composite score, most parcels achieved good performance. For working memory task overall accuracy (*2-back Acc*), the best performing parcels were those in the cingulate cortex, parietal cortex, supramarginal gyrus, lateral frontal cortex and anterior insula. For working memory task performance specifically for faces (*2-back Acc Face*), the distribution was similar to that of the working memory task overall accuracy; in addition, parcels in the inferior temporal cortex and the calcarine sulcus achieved good performance in predicting working memory abilities for faces.

Figure 5e and 5f shows the HCP group activation maps for working memory task (*tfMRI_WM_2BK-0BK*) and social task (*tfmri_SOCIAL_TOM-RANDOM*), where activations are shown in absolute values. Overall, the pattern in the prediction accuracy distribution for

working memory overall accuracy is more similar to the working memory activation pattern, than with the social task activation pattern. In particular, high signal changes in working memory task and high prediction accuracies were both found in the anterior cingulate cortex, parietal cortex, lateral frontal cortex and anterior insula. In contrast, high signal changes in social task was found across the ventral network including the temporal lobe, where prediction accuracies were generally low for working memory performance.

Figures S4 to S7 shows the prediction accuracy distribution using FIX-Pearson-SVR combinations across all 4 granularities, for the 4 selected psychometric variables respectively. The overall distribution patterns were similar across different granularities. Nevertheless, within broad functional territories, the relatively higher predictive power of specific parcels (hence subregions) appeared when reaching granularities such as 200-parcels and 300-parcels, which could reflect relatively more specific brain region-behavior relationships captured at higher granularities.

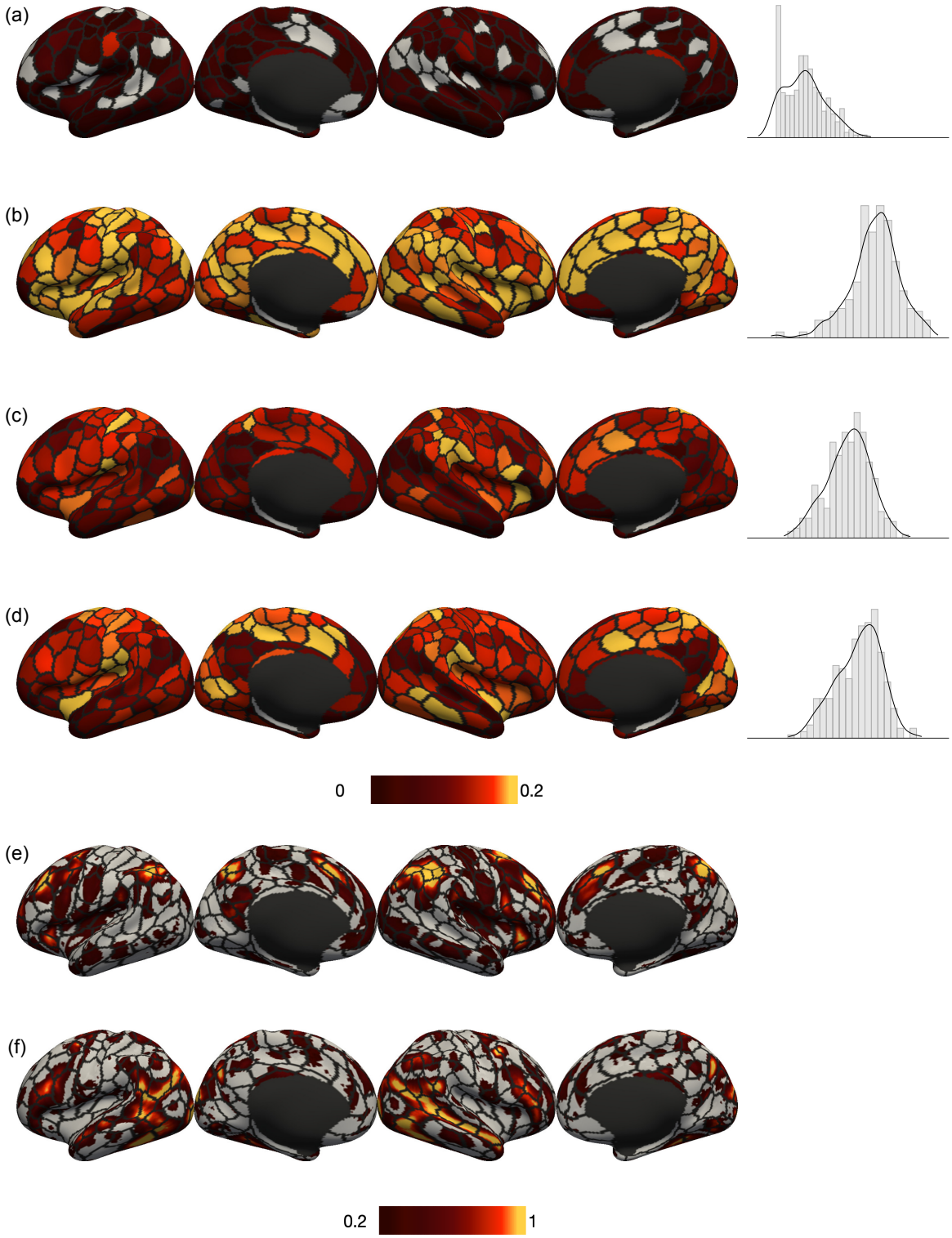

Figure 5. **Top**: Prediction accuracy distribution using FIX-Pearson-SVR combination at 300-parcel granularity of the 4 selected psychometric variables: (a) strength (Strength) (b) crystallized cognition composite score (Crystal Comp) (c) working memory task overall accuracy (2-back Acc) (d) working memory task face condition accuracy (2-back Acc Face). Negative accuracies (i.e. correlation between predicted and actual values) were set to zeros and shown in gray. The rightmost column shows the histogram of prediction accuracies for each variable respectively. **Bottom**: Absolute values of HCP Group activation maps of (e) working memory task (2-back – 0-back) (f) social task (theory of mind – random), with overlay of Schaefer atlas at 300-parcel granularity. The maps were thresholded to show value in the range between 0.2 and 1 (roughly the 50th and 99th percentile respectively). Vertices outside the threshold range were shown in gray.

Effect of denoising on parcel-wise CBPP. To explore the influence of denoising on the validity of the brain-behavior associations, we compared parcel-wise prediction accuracy between minimal preprocessing and FIX denoising data. Figure 6 shows the prediction accuracy distribution and histogram of parcel-wise CBPP across the brain for 4 selected psychometric variables, using minimal-Pearson-SVR combination at 300-parcel granularity. Overall, the prediction accuracies were lower compared to those using FIX data. Furthermore, the variance of prediction accuracies across parcels were significantly lower for minimally processed data than for FIX data, for the strength variable (Levene's test $p < 0.01$; Levene 1960), suggesting that parcels were more differentiated after FIX denoising

To further investigate if FIX denoising contributes to the differentiation between parcels' psychometric profiles or psychometric tests' prediction accuracy distribution maps, we computed the Euclidean distance between 1) the psychometric profiles and 2) accuracy distribution maps for minimally processed data and for FIX data separately. In both cases, the Euclidean distances were significantly larger following FIX denoising than following minimal processing ($p < 0.0001$), suggesting that both psychometric profiles of different parcels and prediction accuracy distributions for different psychometric variables were more dissimilar when FIX denoising was used.

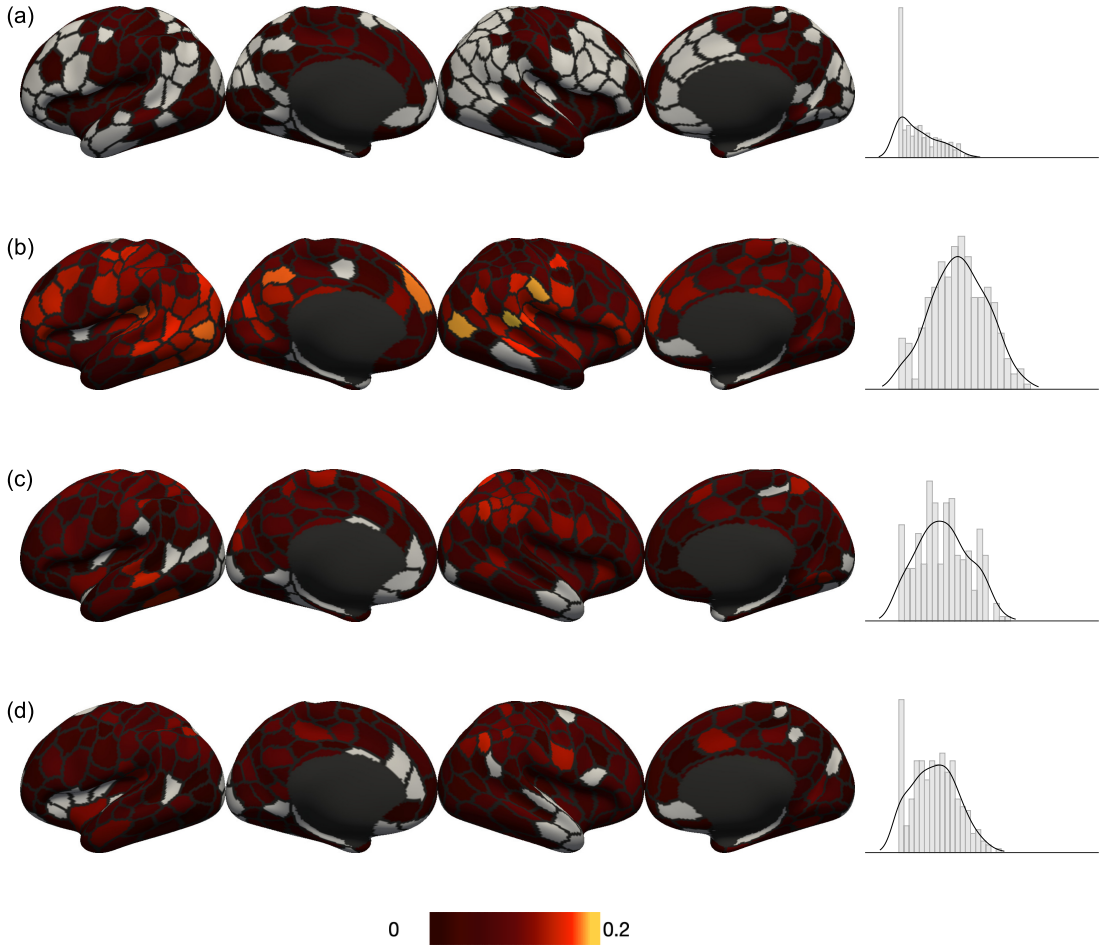

Figure 6. Prediction accuracy distribution using minimal-Pearson-SVR combination at 300-parcel granularity of the 4 selected psychometric variables: (a) strength (Strength) (b) crystallized cognition composite score (Crystal Comp) (c) working memory task overall accuracy (2-back Acc) (d) working memory task face condition accuracy (2-back Acc Face). Negative accuracies (i.e. correlation between predicted and actual values) were set to zeros and shown in gray. The rightmost column shows the histogram of prediction accuracies for each variable respectively.

Effect of confounds. To further investigate the impact of confounds, we implemented the combination of FIX+GSR data followed by SVR using 'no confound' approach, 'confounds added as features' approach, and 'sex + brain size confounds' approach. In the 'no confound' approach, the confound controlling step was completely skipped. In the 'confounds added as features' approach, confounding variables were normalized to zero mean and unit variance, and then appended to the feature matrix used for prediction; these confounding variables were not regressed out in this case either. In the preliminary analysis, when refraining from confounds removal, we found strikingly high prediction accuracy for strength, with further evidence that sex, brain size and intracranial volume were highly correlated with strength (Pearson r = 0.77, 0.54, 0.51 respectively). To further investigate whether these confounds account for the inflated prediction performance, in the 'sex + brain size confounds' approach, only these confounds were regressed out.

Figure 7 shows the whole-brain CBPP results of the 40 selected psychometric variables following the different confound controlling approaches. On one hand, not regressing out confounds or using confounds as features appeared to improve prediction accuracy overall. On the other hand, these two approaches showed extremely high prediction accuracy for the variable 'strength', a measure that would be expected, to a great extent, to be related to demographic and physical characteristics. Our standard regression approach appeared to greatly neutralize such outstanding prediction performance, while simply regressing out sex and brain volume-related confounds had a similar effect.

Figure S8 shows that such a pattern is replicated at the parcel specific prediction level (see Parcel-wise CBPP in Methods), by depicting the prediction results for pairs of parcels in the supramarginal gyrus and Broca region respectively. Similarly, extremely high prediction accuracies for strength were observed, even though both regions are mainly cognition-related. This suggests that not controlling for confounding variables (such as brain size) could undermine the interpretability of parcel-wise CBPP.

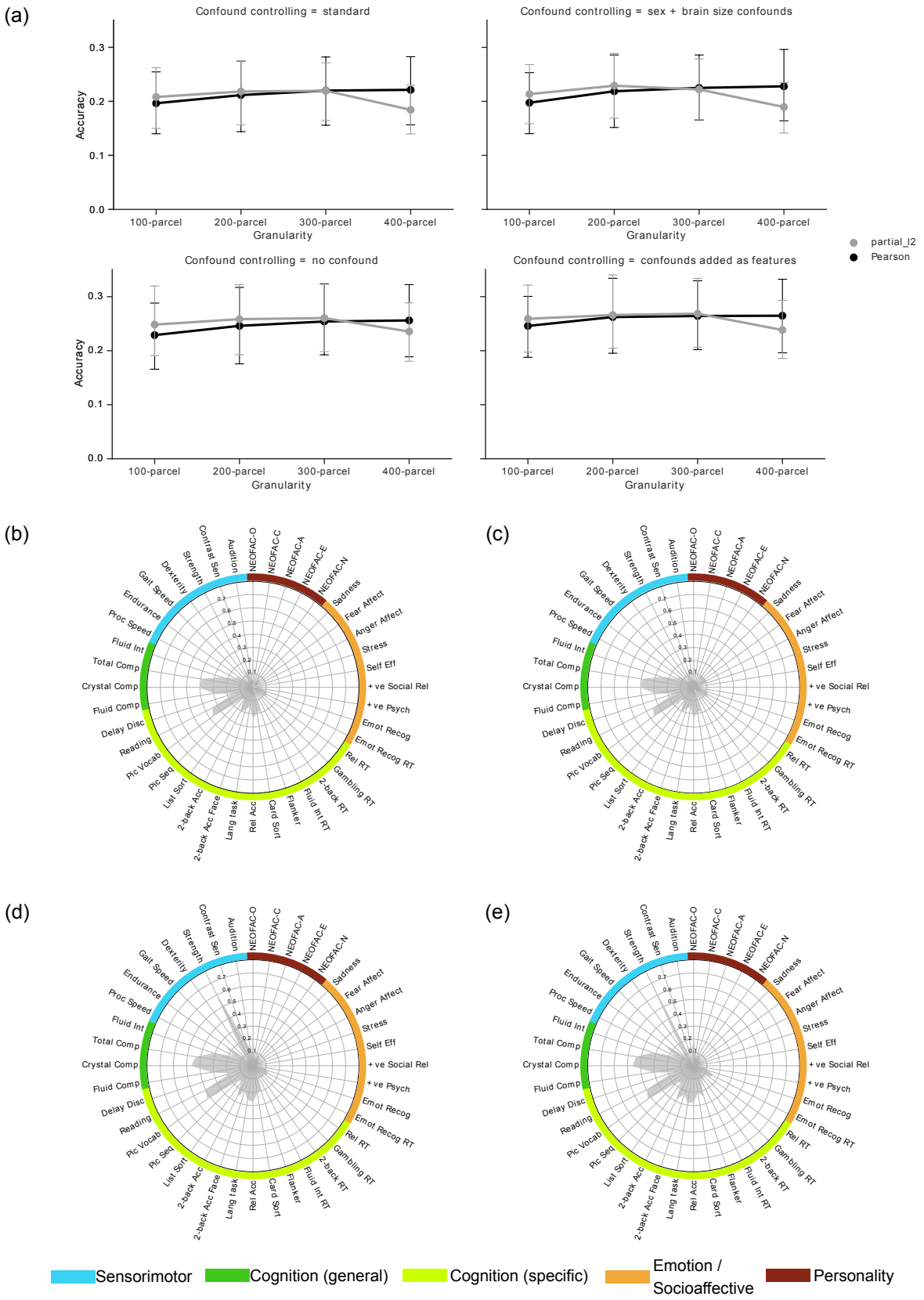

Figure 7. prediction accuracy of (a) top 20 psychometric variables for the different confound controlling approaches across granularity and 40 selected psychometric variables for (b) 'standard' (c) 'sex + brain size confounds' (d) 'no confound' (e) 'confounds added as features' confound controlling approaches at 300-parcel granularity.

Discussion

To develop an optimal framework of connectivity-based prediction of psychometric variables (CBPP) for cognitive neuroscience, we first evaluated the effects of different approaches and parameters that can be used across the different steps of a CBPP. We hence probed the effect of preprocessing, parcellation granularity, FC estimation methods, machine learning algorithm and confounds removal on whole-brain CBPP approaches using 72 different combinations of approaches. Overall, our results demonstrated the relevance of sophisticated denoising approaches for resting fMRI data. Our work also highlighted the importance of confounds removal for the validity of CBPP in a cognitive neuroscience perspective, as well as the good performance of standard regression-based prediction algorithms. Below we discuss how this extensive evaluation combined to an innovative region-wise machine learning approach offer a framework that optimizes both, prediction performance and neurobiological validity (and hence interpretability) of the model.

At the pure quantitative level, the best performing combinations involve using partial correlation and SVR at 200-parcel or 300-parcel granularity, with any preprocessing approach. Nevertheless, if Pearson correlation would be preferred for interpretability purpose, the optimal combination could use both SVR and EN at 200-parcel or 300-parcel granularity, with FIX+GSR preprocessing. At this quantitative level, we should note that most prediction accuracies in our results are in the range of $r = 0.1$ to 0.3. These values, while in line with previous literature, point out to the fact that, even when using whole brain connectivity information, the part of variance across healthy individuals in psychological tests performance that we can explain by individual differences in FC remains relatively low. This issue obviously limits the use of functional connectivity-based prediction in clinical settings, for which strategies like adding confounding variables to features can be utilized to improve prediction performance. In the current study, we focused on the relative prediction accuracies across different parcels or psychometric variables as an insight into the validity of the machine learning approach and, hence, in order to promote interpretability of the framework in a cognitive neuroscience perspective.

Our extensive combination of approaches reveals how various features processing such as denoising, compression and connectivity computation interact and influence the prediction at the quantitative and qualitative level. The broader view made possible by these combinations in turn gave us insight into the possible mechanisms by which features transformations influence the prediction. First our findings show the benefit of sophisticated machine learning-based denoising (here FIX denoising), not only at the quantitative level on prediction performance but also at the qualitative level for the study of brain-behavior relationships. We can note however that, while prediction accuracies were significantly lower for minimally processed data in comparison to FIX and FIX+GSR data when Pearson correlation was used, the three different preprocessing approaches were comparable when partial correlation was used. Considering that the computation of partial correlation between two parcels, to a certain extent, implies the neutralization of common variance, we here assume that using partial correlation potentially removes noise and artifacts affecting several parcels. As a result, similar denoising effects could be achieved by computing partial correlation and by FIX (or FIX+GSR)

preprocessing (also see figure S9). Nevertheless, our supplementary investigation suggests that partial correlation aggressively alters the complex pattern of connectivity fingerprint of brain region (i.e. parcels) by removing shared variance across regions. In line with this view, our preliminary analysis (not reported here) revealed that the residual connectivity fingerprint of the region following partial correlation does not lead to distinguishable psychometric profiles for different parcels. Consequently, these combinations were excluded from the final analysis. At the current stage of knowledge, we can only assume that the common variance removed in partial correlation computation contains variance of interest (e.g. common variance of parcels in the same functional network) that importantly contributes to the region's FC fingerprint. Therefore, while using partial correlation may improve overall prediction performance, it may not provide the same level of interpretability as models using Pearson correlation gives when using region-wise predictions.

Regarding the impact of data compression, prediction accuracies appeared to increase with parcellation granularity. This general trend could be expected since finer parcellation schemes offers more features for prediction, as well as higher functional homogeneity within parcels. However, our combinations using partial correlation revealed a decrease in prediction accuracy at very high level of subdivision (400 parcels). We here assume that this effect is related to the fact that parcels defined based on a group-level parcellation become less accurate with regards to individual brain organization and thus less functionally distinct at this granularity (i.e. some parcels/regions would not fit to specific functional subunits in some subjects). Accordingly, using partial correlation based on a group parcellation at high granularity would not be recommended.

Comparing the regression methods, we showed that a simple implementation of linear regression methods with SVR could perform as well as a more complex implementation with EN, even outperforming the latter when partial correlation was used. Nevertheless, EN would show usefulness in a design where feature weights assignments are to be investigated but partial correlation cannot be used, since EN enforces sparsity in feature weights. When using the current parcel-wise CBPP framework, where interpretability is focused on individual parcel's prediction power rather than the weight of connection between parcels, SVR would be advised.

In terms of confounding effects, on average, removing the effects of demographical, anthropometric and acquisition factors lead to lower prediction accuracies. However, our results showed that this is mainly due to a large decrease in prediction power for strength performance. The higher strength prediction power observed when keeping variance related to confounding factors or when adding these confounding factors as features in the predictive models could be attributed to the confounding effects of head size and sex. Our investigations hence demonstrate that both factors can compromise the validity of the results when the objective is to relate brain functional connectivity to behavioral measures including the evaluation of the predictive power of individual brain regions' connectivity for psychometric variables.

Capitalizing on our evaluation of approaches using whole brain information (i.e. whole brain connectivity features), we then implemented a region-wise (i.e. parcel-wise) prediction approach and qualitatively examined the psychometric prediction profile for each parcel, which

overall matched well our expectations based on brain mapping literature. Hence, the parcels in primary visual cortex showed predictive power for processing speed performance. This pattern could be explained by the role of these regions in visual information processing and relaying (Fabre-Thorpe et al. 2001; Sharpee et al. 2006). The performance at two higher cognitive variables (executive function type: *Card Sort* and *Flanker*) were also relatively well predicted from primary visual cortex's FC patterns, which could potentially be explained by the role of visual processing speed and accuracy in the performance. The parcels in supramarginal gyrus showed predictive power across a wide range of higher cognitive function measures, but also for emotion recognition tasks, in line with the brain mapping literature suggesting the involvement of the temporoparietal junction in social processes, such as empathy (Lawrence et al. 2006; Schnell et al. 2011; Carlson 2012). Additionally, overall, the FC patterns of the parcels in supramarginal gyrus and the Broca region show predictive power for psychometric variables related to language processing and crystallized intelligence. Both regions have been previously shown to be involved in word and phonological processing (Binder et al. 1997; McNealy et al. 2006; Klepousniotou et al. 2014; Twomey et al. 2015; Oberhuber et al. 2016). Comparing the prediction profile of the left and right Broca region further supports the validity of our approach since, in contrast to the left Broca region, the right Broca region showed relatively low predictive power for language task performance, a pattern that is in agreement with the left hemisphere's dominance for language function (Clos et al. 2013; Karsolis et al. 2019). Furthermore, parcels in these two regions also showed predictive power for working memory measures, in agreement with their role in the phonological loop in traditional working memory models (Salmon et al. 1996; Smith and Jonides 1999; Zurowski et al. 2002; Rogalsky and Hickok 2011).

Our examination of the psychometric prediction accuracy distribution maps appeared to converge also well with the brain mapping literature. When examining specific measures such as strength performance, our prediction maps revealed a relatively specific pattern reminiscent of the sensorimotor network. In contrast, multidetermined score such as the crystallized intelligence composite score revealed a very broad pattern, in which the parcels with the highest prediction power lie in regions that frequently show high interindividual variability in RSFC patterns, such as the supramarginal gyrus and the anterior insula (Mueller et al., 2013; Laumann et al., 2015; Kong et al., 2018). Comparing the pattern of prediction performance for the total accuracy score for the 2-back task and the same score only for faces material revealed overall similar patterns but with higher prediction power in the right hemisphere for faces, in particular in the ventral temporal regions, in agreement with the literature on faces processing in the brain (Sams et al., 1997; Nakamura et al., 2000; Nelson, 2001).

When examining the accuracy distribution maps across granularities level (i.e. region subdivision level), the overall patterns remain similar for every selected psychometric measure. With increased granularity, especially at 200-parcel and 300-parcel granularity, the proportion of parcels showing high predictive power increased. This coincides with our observations of achieving higher prediction performance in whole-brain CBPP with higher granularity, thus reinforcing the conclusion that 200 to 300 parcels represents an optimal range for fMRI surface data. Finally, our observation that the subdivision of broad functional region into specific

parcels at higher granularity goes along with the higher predictive power for some parcels may reflect the fact that these parcels represent functional sub-units, and hence that 200/300 parcels granularity provides a better representation of brain functional data from a cognitive neuroscience standpoint.

In sum, by using high quality data, including a broad range of psychometric measures in a healthy adult cohort, we developed a region-wise connectivity based psychometric prediction framework based on supervised learning approaches linking resting-state functional connectivity of brain regions to behavior. Evaluating the CBPP framework both in terms of prediction performance and interpretability, we demonstrated the relevance of sophisticated denoising approaches for both aspects. Our work also highlighted the importance of confounds removal for the validity of CBPP in a cognitive neuroscience perspective, as well as the good performance of simple regression-based prediction algorithm. To promote the use of our specific region-wise approach in cognitive neuroscience studies, we illustrated two main applications for which we evaluated the brain region-behavior relationships: 1) psychometric profiles of brain regions; 2) brain maps of prediction accuracies distribution for specific psychometric variables. The material to implement our approach is openly available at https://github.com/inm7/cbpp.The developed framework could also contribute to a better understanding of the relationships between brain regions' connectivity and behavioral phenotypes in aging and clinical populations. The transferability of our framework in this context should be investigated in future studies.

Methods

Data and preprocessing. For our evaluation study, the 1200 Subjects Data Release of the HCP data was used (Van Essen et al. 2013). Subjects were young healthy adults (aged 22-37), from families with twins and non-twin siblings. Imaging data were acquired using a customized Siemens 3T Skyra. Each subject visited in two consecutive days, during each of which two rs-fMRI runs were acquired using different phase encodings, left-right and right-left. The subjects were asked to (relaxingly) fixate on a projected bright cross-hair on a dark background during the run. Each run is 1200 frames (14.4 min) in length, with a repetition time (TR) of 720 ms.

We used the readily available preprocessed data from the HCP initiative. All the raw rs-fMRI data from HCP were preprocessed by the HCP Minimal Processing Pipelines (Glasser et al. 2013). We refer to the resulting data as "minimally processed". Further cleaning of noise in HCP data was done using ICA-FIX. Components were first identified using independent component analysis (ICA); a classifier (FIX) trained on HCP data was then applied to remove artifactual components (Smith et al. 2013; Salimi-Khorshidi et al. 2014; Griffanti et al. 2014). We refer to these data as "FIX". To further investigate the impact of global signal regression (GSR), we regressed the cortical global signal, which is the average signal across all cortical vertices, and its temporal derivative (Power et al. 2018; Li et al. 2019) from the FIX data. These data are then referred to as "FIX+GSR".

For most subjects of the HCP cohort, a large number of psychological measures were collected through tests and questionnaires across a broad range of psychological domains including sensory, motor, cognition, emotion, affect and personality. Additionally, task performance scores were also available from task-fMRI sessions. We selected 98 psychometric variables after inspecting all variables distribution (see Table S1) for roughly normal distribution.

Psychometric data are typically related to demographical factors such as age and gender. In addition, the relationships between psychometric and neuroimaging data can be partially mediated by factors such as head size, handedness and seasonality. To ensure strict control on the confounding effects of these latter variables, we not only regressed out the first level effect of brain size, sex and age but also included the secondary term of age and the interaction between age and sex, in line with the HCP MegaTrawl analysis (Smith et al. 2016). Accordingly, our 'standard' confound controlling method involved regressing out 9 confounding variables from the psychometric variables: sex (*Gender*), age (*Age_in_Yrs*), age^2, sex*age, sex*age^2, handedness (*Handedness*), brain size (*FS_BrainSeg_Vol*), intracranial volume (ICV; *FS_IntraCranial_Vol*) and acquisition quarter (*Acquisition*).

Data compression. CBPP studies usually capitalizes on whole-brain connectivity using atlases (brain parcellation) to summarize voxels data, by computing a mean time series for each parcel by averaging time series of all the vertices (or voxels) within that parcel. In our implementation, the Schaefer atlas (Schaefer et al. 2018) was utilized at 4 different granularity levels, with 100 parcels, 200 parcels, 300 parcels and 400 parcels respectively. The most commonly used functional atlases contain 200 to 400 parcels (Shen et al. 2013; Joliot et al. 2015; Gordon et al. 2016; Glasser et al. 2016). Nonetheless, in the recent HCP MegaTrawl, the highest psychometric prediction accuracy was achieved at lower dimensionality, with dimension d = 50, when partial correlation (with L2 regularization) was used for computing connectivity (Smith et al. 2016). Therefore, we used all available granularity up to 400-parcel level.

Functional connectivity computation. Functional connectivity is typically computed based on Pearson correlation; however, partial correlation is considered by several authors as more likely to reflect direct connections (Bijsterbosch et al. 2017; Lim et al. 2019). When using atlases, functional connectivity strengths are computed as the correlation coefficient between each pair of parcel's average time series. In our implementation, the Pearson correlation matrices were computed using the Matlab's corrcoef function. The partial correlation matrices were computed using FSLNets' nets_netmats function (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLNets), with the default regularization coefficient for L2-norm ridge regression. Since 4 runs were acquired for each HCP subject, the average connectivity matrix for each subject across runs was used as input features for the supervised learning model.

Psychometric prediction. For the prediction step, linear SVR was implemented using Matlab's fitrlinear, while EN was implemented using the glmnet package for Matlab (Qian et al. 2013). We note that ridge regression and lasso are also widely used in neuroscience studies. However, since EN was used in previous CBPP studies (Smith et al. 2016; Dubois et al. 2018; Kashyap et al. 2019) and could be considered as an optimal combination of ridge regression and lasso, harnessing both the high prediction performance and sparse representation (Zou and Hastie 2005), we only included EN.

Importantly, feature selection can be done before the prediction step to reduce computation time (Smith et al. 2016; Shen et al. 2017; Dubois et al. 2018 Kashyap et al. 2019). One popular way to do so is to examine Pearson correlation coefficients between elements of the connectivity matrix (i.e. edges) and the psychometric variables to predict and to select either connectivity elements with significant correlation ($p < 0.01$ or $p < 0.05$) or the top 50% of all the connectivity elements (i.e. 2475, 9950, 22425 and 39900 elements at 100-parcel, 200-parce, 300-parcel, and 400-parcel granularity respectively) as features for the subsequent prediction step. In our implementation, we selected the top 50% of all connectivity elements as features for the more computationally expensive algorithm, EN. As MLR relies on least square solutions and was mostly used in previous studies for univariate prediction, it could not be expected to deal with higher feature dimensionality. Therefore, we only selected the top 500 of all connectivity elements as features for MLR in order to prevent overfitting. No feature selection was performed for linear SVR since the computation was not particularly expensive and, importantly, no overfitting was observed.

Finally, psychometric prediction is typically performed by training and testing regression algorithms through cross-validation. For each combination of approaches, we performed 10 repeats of 10-fold cross-validation. As the HCP cohort consists of families of siblings, the cross-validation folds were generated when ensuring that family members were always kept within the same fold, as done in previous studies (Kong et al. 2018; Li et al. 2019). Using Matlab's regress function, the 9 confounding variables were regressed out from the training set from 9 folds; the same regression coefficients were used to remove effects of these variables from the test set in the remaining fold. Similarly, feature selection was performed based on the training data only; the same selected features were used in the test fold. The difference in performance between combinations were tested with the corrected resampled t-test (Nadeau and Bengio 2003; Bouckaert and Frank 2004), which accounts for the dependency between test folds across repeats. The results were corrected for multiple comparisons using false discovery rate (FDR; Benjamini & Hochberg 1995) of $q < 0.05$.

The hyperparameter, epsilon, for linear SVR was set to the default value. As the default value is dependent on data variance and usually found to be the most optimal value during our preliminary analysis, our final implementation of linear SVR did not include hyperparameter tuning. Overall, our implementation of linear SVR can be seen as a simpler linear regression model, in comparison to EN which would rather require hyperparameter tuning. Hyperparameter tuning for EN was done in two steps. During each cross-validation run, we first fixed the alpha value, which determines the compromise between ridge and lasso, while tuning the lambda value, which determines the degree of regularization. To do so, a 10-fold

inner-loop cross-validation was carried out using 8 of the training folds. The best alpha value was then chosen according to the prediction performance on the last training fold.

Parcel-wise CBPP. To investigate the neurobiological validity and hence interpretability of CBPP results, we performed parcel-wise CBPP using individual parcel's connectivity profile. A parcel's connectivity profile is represented by a vector of functional connectivity values between the parcel and all other parcels. No feature selection based on correlation with the psychometric variables was performed for parcel-wise prediction. Permutation testing was performed by repeating 10-fold cross-validation 1000 times while shuffling the psychometric variables. The final prediction accuracies across the 98 psychometric variables were corrected for multiple comparisons using false discovery rate (FDR; Benjamini & Hochberg 1995) of $q < 0.05$. For generating the prediction accuracies distributions of psychometric variables, permutation testing was not performed due to the very larger number of parcels involved (especially at higher granularities).

To facilitate the inspection of prediction profiles, the variables were summarized into 40 variables or combined variables, mostly by opting for age-adjusted scores over unadjusted ones, as well as general scores over sub-category scores, and finally by combining some sub-category scores to a more general score. In this latter case, the prediction accuracy was obtained as the average across all the combined scores. See Table S2 for the complete list of variables selection.

Code availability

All codes are publicly available at https://github.com/inm7/cbpp.

References

Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T. 1997. Human brain language areas indentified by functional Magnetic Resonance Imaging. Journal of Neuroscience. 17(1): 353-362.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Socity. 57:289-300.

Bijsterbosch J, Smith SM, Beckmann CF. 2017. Introduction to resting state fMRI functional connectivity. Oxford University Press.

Boser EB, Guyon IM, Vapnik VN. 1992. A training algorithm for optimal margin classifiers. Proc fifth annual workshop on COTL. 144-152.

Bouckaert RR, Frank E. 2004. Evaluating the replicability of significance tests for comparing learning algorithms. Adv. Knowl. Discov. Data Min. 3-12.

Caspers S, Moebus S, Lux S, Pundt N, Schütz H, Mühleisen TW, Gras V, Eickhoff SB, Romanzetti S, Stöcker T, et al. 2014. Studying variability in human brain aging in a population-based German cohort – Rationale and design of 1000BRAINS. Front Aging Neurosci. 6:149.

Carlson NR. 2012. Physiology of behavior. 11th edition. Pearson.

Chatterjee S, Hadi AS. 1986. Influential observations, high leverage points, and outliers in linear regression. Stat Sci. 1:379-416.

Clos M, Amunts K, Laird AR, Fox PT, Eickhoff SB. 2013. Tackling the multifunctional nature of Broca's region meta-analytically: Co-activation-based parcellation of area 44. NeuroImage. 83:174-188.

Cortes C, Vapnik VN. 1995. Support-vector networks. Machine Learning. 20(3):273-297.

Dadi K, Rahim M, Abraham A, Chyzhyk D, Milham M, Thirion B, Varoquaux G, for the Alzheimer's Disease Neuroimaging Initiative. 2019. Benchmarking functional connectome-based predictive models for resting-state fMRI. NeuroImage. 192:115-134.

Dubois J, Galdi P, Han Y-T, Paul LK, Adolphs R. 2018. Resting-state functional brain connectivity best predicts personality dimension of openness to experience. Personality Neuroscience. 1(e6):1-21.

Fabre-Thorpe M, Delorme A, Marlot C, Thorpe S. 2001. A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. Journal of Cognitive Neuroscience. 13(2):171-180.

Finn ES, Shen X-L, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetric X, Constable RT. 2015. Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. Nature Neuroscience. 18(11):1664-1671.

Griffanti L, Salimi-Khorshidi G, Beckmann CF, Auerbach EJ, Douaud G, Sexton CE, Zsoldos E, Ebmeier KP, Filippini N, Mackay CE, et al. 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. NeuroImage. 95:232-247.

Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J-Q, Jbabdi S, Webster M, Polimeni JR, et al. 2013. The minimal preprocessing pipeline for the Human Connectome Project. NeuroImage. 80:105-124.

Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, et al. 2016. A multi-modal parcellation of human cerebral cortex. Nature. 536(7615):171-178.

Gordon EM, Laumann TO, Adeyemo B, Huckins JK, Kelley WM, Petersen SE. 2014. Generation and evaluation of a cortical area parcellation from resting-state correlations. Cerebral Cortex. 26(1):288-303.

Joliot M, Jobard G, Naveau M, Delcroix N, Petit L, Zago L, Crivello F, Mellet E, Mazoyer B, Tzourio-Mazoyer N. 2015. AICHA: An atlas of intrinsic connectivity of homotopic areas. Journal of Neuroscience Methods. 254:46-59.

Karolis VR, Corbetta M, de Schotten MT. 2019. The architecture of functional lateralisation and its relationships to callosal connectivity in the human brain. Nature Communications. 10(1):1417.

Kashyap R, Kong R, Bhattacharjee S, Li J-W, Zhou J, Yeo BTT. 2019. Individual-specific fMRI-subspaces improve functional connectivity prediction of behavior. NeuroImage. 189:804-812.

Klepousniotou E, Gracco VL, Pike GB. 2014. Pathways to lexical ambiguity: fMRI evidence for bilateral fronto-parietal involvement in language processing. Brain Lang. 131:56-64.

Kong R, Li J-W, Orban C, Sabuncu MR, Liu H-S, Schaefer A, Sun N-B, Zuo X-N, Holmes AJ, Eickhoff SB, Yeo BTT. 2018. Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion. Cerebral Cortex. 29:2533-2551.

Laumann TO, Gordon EM, Adeyemo B, Snyder AZ, Joo SJ, Chen M-Y, Gilmore AW, McDermott KB, Nelson SM, Dosenbach NUF, et al. 2015. Functional system and areal organization of a highly sampled individual human brain. Neuron. 87(3):657-670.

Lawrence EJ, Shaw P, Giampietro VP, Surguladze S, Brammer MJ, David AS. 2006. The role of 'shared representations' in social perception and empathy: An fMRI study. NeuroImage. 29:1173-1184.

Levene H. 1960. Robust tests for equality of variances. In: Olkin I, Hotelling H, editor. Contributions to probability and statistics: Essays in honor of Harold Hotelling. Stanford University Press. p 278-292.

Li J-W, Kong R, Liegeois R, Orban C, Tan Y-R, Sun N-B, Holmes AJ, Sabuncu MR, Ge T, Yeo BTT. 2019. Global signal regression strengthens association between resting-state functional connectivity and behaviour. NeuroImage. 196:126-141.

Lim S, Radicchi F, van den Heuvel MP, Sporns O. 2019. Discordant attributes of structural and functional brain connectivity in a two-layer multiplex network. Scientific Reports. 9:2885.

Matsui T, Murakami T, Ohki K. 2016. Transient neuronal coactivations embedded in globally propagating waves underlie resting-state functional connectivity. Proc Nat Acad Sci. 113:6556-6561.

McNealy K, Mazziotta JC, Dapretto M. 2006. Cracking the language code: neural mechanisms underlying speech parsing. Journal of Neuroscience. 26(29): 7629-7638.

Mueller S, Wang D-H, Fox MD, Yeo BTT, Sepulcre J, Sabuncu MR, Shafee R, Liu J, Liu H-S. 2013. Individual variability in functional connectivity architecture of the human brain. Neuron. 77(3):586-595.

Nakamura K, Kawashima R, Sato N, Nakamura A, Sugiura M, Kato T, Hatano K, Ito K, Fukuda H, Schormann T, Zilles K. 2000. Functional delineation of the human occipito-temporal areas related to face and scene processing: A PET study. Brain. 123(9):1903-1912.

Naudeau C, Bengio Y. 2003. Inference for the generalization error. Mach. Learn. 52:239-281.

Nelson CA. 2001. The development and neural bases of face recognition. Inf Child Dev. 10:3-18.

Noble S, Spann MN, Tokoglu F, Shen X-L, Constable RT, Scheinost D. 2017. Influences on the test-retest reliability of functional connectivity MRI and its relationship with behavioural utility. Cerebral Cortex. 27:5415-5429.

Oberhuber M, Hope TM, Seghier ML, Parker JO, Prejawa S, Green DW, Price CJ. 2016. Four functionally distinct regions in the left supramarginal gyrus support word processing. Cerebral Cortex. 26(11): 4212-4226.

Plachti A, Eickhoff SB, Hoffstaedter F, Patil KR, Laird AR, Fox PT, Amunts K, Genon S. 2019. Multimodal parcellations and extensive behavioural profiling tackling the hippocampus gradient. Cerebral Cortex. In Press. doi: 10.1093/cercor/bhy336.

Power JD, Plitt M, Gotts SJ, Kundu P, Voon V, Bandettini PA, Martin A. 2018. Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data. Proc Nat Acad Sci. 115:e2105-2114.

Qian J, Hastie T, Friedman J, Tibshirani R, Simon N. 2013. Glmnet for Matlab. Available at: http://www.stanford.edu/~hastie/glmnet_matlab/ (Accessed: 15 Mar 2019).

Rogalsky C, Hickok G. 2011. The role of Broca's area in sentence comprehension. Journal of Cognitive Neuroscience. 23(7):1664-1680.

Rosenverg MD, Finn ES, Scheinost D, Papademetris X, Shen X-L, Constable RT, Chun MM. 2016. A neuromarker of sustained attention from whole-brain functional connectivity. Nature Neuroscience. 19(1):165-171.

Salimi-Khorshidi G, Dounaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM. 2014. Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. NeuroImage. 90:449-468.

Salmon E, Van der Linden M, Collette F, Delfiore G, Maquet P, Degueldre C, Luxen A, Franck G. 1996. Regional brain activity during working memory tasks. Brain. 119(5):1617-1625.

Sams M, Hietanen JK, Hari R, Ilmoniemi RJ, Lounasmaa OV. 1997. Face-specific responses from the human inferior occipito-temporal cortex. Neuroscience. 71(1):49-55.

Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, Eickhoff SB, Yeo BTT. 2018. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cerebral Cortex. 28:3095-3114.

Scheinost D, Noble S, Horien C, Greene AS, Lake EMR, Mehraveh S, Gao S-Y, Shen X-L, O'Connor D, Barron DS, et al. 2019. Ten simple rules for predictive modelling of individual differences in neuroimaging. NeuroImage. 193:35-45.

Schnell K, Bluschke S, Konradt B, Walter H. 2011. Functional relations of empathy and mentalizing: An fMRI study on the neural basis of cognitive empathy. NeuroImage. 54:1743-1754.

Schölvinck ML, Maier A, Frank QY, Duyn JH, Leopold DA. 2010. Neural basis of global resting-state fMRI activity. Proc Nat Acad Sci. 107:10238-10243.

Sharpee TO, Sugihara H, Kurgansky AV, Rebrik SP, Stryker MP, Miller KD. 2006. Adaptive filtering enhances information transmission in visual cortex. Nature. 439(7090):936-942.

Shen X-L, Tokoglu F, Papademetris X, Constable RT. 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. NeuroImage. 82:403-415.

Shen X-L, Finn ES, Scheinost D, Rosenberg MD, Chun MM, Papademetris X, Constable RT. 2017. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. Nat Protoc. 12:506-518.

Smith EE, Jonides J. 1999. Storage and executive processes in the frontal lobes. Science. 283(5408):1657-1661.

Smith SM, Beckmann CF, Andersson J, Auerbach E, Bijsterbosch J, Douaud G, Duff E, Feinberg DA, Griffanti L, Harms MP, et al. 2013. Resting-state fMRI in the Human Connectome Project. NeuroImage. 80:144-168.

Smith SM, Vidaurre D, Glasser M, Winkler A, McCarthy P, Robinson E, Chen X, Horton W, Jenkinson M, Duff E, et al. 2016. Second beta-release of the HCP functional connectivity MegaTrawl.

Twomey T, Waters D, Price CJ, Kherif F, Woll B, MacSweeney M. 2015. Identification of the regions involved in phonological assembly using a novel paradigm. Brain Lang. 150:45-53.

Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, for the WU-Minn HCP Consortium. 2013. The Wu-Minn Human Connectome Project: An overview. NeuroImage. 80:62-79.

Weis S, Patil KR, Hoffstaedter F, Nostro A, Yeo BTT, Eickhoff SB. 2019. Sex classification by resting state brain connectivity. Cerebral Cortex. In press. doi:10.1093/cercor/bhz129.

Wen H-G, Liu Z-M. 2016. Broadband electrophysiological dynamics contribute to global resting-state fMRI signal. Journal of Neuroscience. 36(22):6030-6040.

Zou H, Hastie T. 2005. Regularization and variable selection via the elastic net. J R Statist Soc B. 67:301-320.

Zurowski B, Gostomzyk J, Grön G, Weller R, Schirrmeister H, Neumeier B, Spitzer M, Reske SV, Walter H. 2002. Dissociating a common working memory network from different neural substrates of phonological and spatial stimulus processing. NeuroImage. 15(1):45-57.

Acknowledgement

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, GE 2835/1–1, EI 816/ 4–1), the Helmholtz Portfolio Theme ‘Supercomputing and Modelling for the Human Brain’ and the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 720270 (HBP SGA1) and Grant Agreement No. 785907 (HBP SGA2). Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The authors also thank Hesheng Liu for helpful discussion.

Author contributions

J.W., K.R.P., and S.G. designed research; J.W. performed research; J.W., F.H., and S.G. analyzed data; and J.W., S.B.E., H.S., and S.G. wrote the paper.

Table 1. Summary of existing psychometric prediction approaches in the framework of CBPP

Reference	Preprocessing	Parcellation granularity & Connectivity	Feature selection	Confounds	Regression
(Fin et al., 2015)	HCP minimal preprocessing pipeline, nuisance regression (motion, WM, CSV, and GS), detrending, LPF	268x268 Pearson correlation matrix	Edges with significant correlation (P<0.01, P<0.05 and P<0.10) with target, summed according to sign (positive/negative)	Nil	Multiple linear regression with LOOCV
(Rosenberg et al., 2016)	MC, nuisance regression (motion, WM, CSF, GM), temporal smoothing	268x268 Pearson correlation matrix	Edges with significant correlation (P<0.01) with target, summed according to sign	Nil	Multiple linear regression with LOOCV
(Smith et al., 2016)	HCP minimal preprocessing pipeline, ICA-FIX	Group-ICA of 15, 25, 50, 100, 200, 300 parcels; partial correlation with L2 regularization	Top 50% edges by correlation with targets	sex, age, age^2, sex*age, sex*age^2, brain & head size, overall head motion, acquisition quarter	Elastic net with leave-one-family-out cross-validation
(Noble et al., 2017)	HCP minimal preprocessing pipeline, nuisance regression (motion, WM, CSF, GS), detrending, LPF	268x268 Pearson correlation matrix	Edges with significant correlation (P<0.05) with target, summed according to sign	Nil	Multiple linear regression with LOOCV
(Dubois et al., 2018)	HCP minimal preprocessing, followed by: i) nuisance regression (motion, WM, CSF), LPF, GSR ii) detrending, BPF, nuisance regression (motion, WM, CSF, GS) combined with censoring iii) ICA-FIX, detrending, nuisance regression (GM, GS, CompCor of WM & CSF), censoring, BPF	i) 268x268 Pearson correlation matrix ii) 360x360 Pearson correlation matrix	i) edges with significant correlation (P<0.01) with target, summed according to sign ii) edges with significant correlation (P<0.01) with target	sex, age, handedness, fluid intelligence, brain size, motion, multiband reconstruction algorithm	Elastic net with leave-one-family-out cross-validation
(Li et al., 2019)	i) GSP data: STC, MC, nuisance regression (motion, WM, CSF, with/without GS), censoring, BPF ii) HCP data: HCP minimal preprocessing, ICA-FIX, censoring, with/without GSR	419x419 Pearson correlation matrix	Nil	sex, age, FD, DVARS	kernel ridge regression with 20-fold cross-validation
(Kashyap et al., 2019)	HCP minimal preprocessing pipeline, ICA-FIX, censoring, nuisance regression (motion, WM, CSF, GS), BPF	i) 419x419 Pearson correlation matrix ii) 419x419 partial correlation matrix	Top 50% edges by correlation with targets	sex, age, FD	Elastic net with 20-fold cross-validation

BPF, band-pass filtering; CSF, cerebrospinal fluid signal; DVARS, a summary score for data quality, defined as the spatial standard deviation of successive difference images; FD, frame-wise displacement; GM, gray matter signal; GS, global signal; GSR, global signal regression; GSP, Brain Genomics Superstruct Project; HCP: Human Connectome Project; LPF, low-pass filtering; MC, motion correction; STC, slice time correction; WM, white matter signal.