From the Institute for System Neuroscience
at Heinrich Heine University Düsseldorf

Brain Region-wise Connectivity-based Psychometric Prediction

Framework, Interpretation, Replicability and Generalizability

Dissertation

to obtain the academic title of Doctor of Philosophy (PhD) in Medical Sciences
from the Faculty of Medicine at Heinrich Heine University Düsseldorf

submitted by
Jianxiao Wu
(2022)

As an inaugural dissertation printed by permission of the
Faculty of Medicine at Heinrich Heine University Düsseldorf

signed:
Dean: Prof. Dr. med. Nikolaj Klöcker
Examiners: Prof. Dr. Sarah Genon, Prof. Dr. Holger Schwender

I am walking a road that has been walked before, but to me it is new. It leads to the place where I was born and raised, which soon now I will discover for the first time.

- William Nicholson, *The Society of Others*

Parts of this work have been published:

Wu, J., Eickhoff, S.B., Hoffstaedter, F., Patil, K.R., Schwender, H., Yeo, B.T.T., and Genon, S. 2021. A Connectivity-Based Psychometric Prediction Framework for Brain-Behavior Relationship Studies. *Cerebral Cortex*. 31(8), 3732–3751. DOI: 10.1093/cercor/bhab044

Wu, J., Li, J., Eickhoff, S.B., Hoffstaedter, F., Hanke, M., Yeo, B.T., and Genon, S. 2022. Cross-cohort replicability and generalizability of connectivity-based psychometric prediction patterns. *Neuroimage*, 119569. DOI: 10.1016/j.neuroimage.2022.119569

Yeung, A.W.K., More, S., Wu, J., and Eickhoff, S.B. 2022. Reporting details of neuroimaging studies on individual traits prediction: a literature survey. *Neuroimage*, 119275. DOI: 10.1016/j.neuroimage.2022.119275

Zusammenfassung

Die Untersuchung des Zusammenhangs zwischen Gehirn und Verhalten ist ein essentieller Aspekt der Neurowissenschaften. Die interindividuelle Variabilität psychometrischer Messwerte mit der unterindividuellen Variabilität von Gehirnbilddaten in Beziehung zu setzen wird eine vor allem in letzter Zeit immer beliebtere Untersuchungsmethode. Insbesondere Vorhersageansätze mit Kreuzvalidierung können nützlich sein, um verallgemeinerbare Beziehungen zwischen Gehirn und Verhalten auf datengesteuerte Weise zu identifizieren. Dennoch bleibt zu klären, welche Gehirn-Verhaltens-Beziehungen aus den Vorhersagemodellen interpretiert werden können, sowie die Verallgemeinerbarkeit der Modelle auf völlig neue Kohorten. In dieser Arbeit versuchen wir, die Lücke der Interpretierbarkeit zu schließen, indem wir einen auf regionaler Konnektivität basierenden psychometrischen Vorhersageansatz entwickeln. Dies umfasst ein regionenbezogenes Vorgehen, bei dem für jede Hirnregion ein Vorhersagemodell geschätzt und bewertet wird. Die Vorhersagegenauigkeit jedes regionenspezifischen Modells ist ein direkter Hinweis auf die Assoziation dieser Gehirnregion mit der vorhergesagten psychometrischen Messung. In Studie 1 haben wir dese Vorgehensweise auf eine Reihe psychometrischer Variablen aus einer großen gesunden Kohorte angewendet. Wir konnten zeigen, dass der Ansatz hilfreich ist, um sowohl regionenspezifische psychometrische Vorhersageprofile als auch psychometrische Vorhersagemuster für das gesamte Gehirn zu erstellen. In Studie 2 haben wir die Nützlichkeit des Ansatzes bei der Bewertung der kohortenübergreifenden Replizierbarkeit und Generalisierbarkeit in Bezug auf die aus den Vorhersagemodellen abgeleiteten Beziehungen zwischen Gehirn und Verhalten gezeigt, anstatt nur auf der Grundlage der Vorhersagegenauigkeit. In Studie 3 untersuchten wir systematisch bestehende psychometrische Vorhersagestudien und fassten die Trends auf diesem Gebiet zusammen, welche die Verwendung großer Kohorten und eine externe Validierung forderten. Insgesamt hat unsere Arbeit gezeigt, wie wichtig Interpretierbarkeit und Verallgemeinerbarkeit für psychometrische Vorhersagen sind. Wir empfehlen die Verwendung mehrerer großer Kohorten zur Bewertung der Interpretierbarkeit und Verallgemeinerbarkeit.

Summary

The study of brain-behavior relationships is a fundamental aspect of neuroscience. Recently, it has become increasingly popular to investigate brain-behavior relationships by relating the interindividual variability in psychometric measure to the interindividual variability in brain imaging data. In particular, prediction approaches with cross-validation can be useful for identifying generalizable brain-behavior relationships in a data-driven manner. Nevertheless, it remains to be ascertained what brain-behavior relationships can be interpreted from the prediction models, and how generalizable the models are to fully new cohorts. In this work, we attempt to fill in the gap of interpretability by developing a region-wise connectivity-based psychometric prediction (CBPP) framework. This framework involves a region-wise approach where a prediction model is estimated and evaluated for each brain region. The prediction accuracy of each region-wise model is a direct indication of that brain region's association with the psychometric measure predicted. In study 1, we applied the framework to a range of psychometric variables from a large healthy cohort and demonstrated the helpfulness of the framework in constructing region-wise psychometric prediction profiles or psychometric-wise prediction pattern across the brain. In study 2, we demonstrated the usefulness of the framework in assessing cross-cohort replicability and generalizability in terms of brain-behavior relationships derived from the prediction models, instead of just based on prediction accuracies. In study 3, we systematically examined existing psychometric prediction studies, summarizing the trends in the field, calling for the use of large cohorts and external validation. Overall, our work suggested the importance of interpretability and generalizability for psychometric prediction, recommending the use of multiple large cohorts in evaluating the interpretability and generalizability.

List of abbreviations

BOLD	blood-oxygen-level dependent
CBPP	connectivity-based psychometric prediction
EEG	electroencephalography
EN	elastic net
FC	functional connectivity
fMRI	functional Magnetic Resonance Imaging
ICA	Independent Component Analysis
ICA-AROMA	ICA-based Automatic Removal of Motion Artifacts
ICA-FIX	FMRIB's ICA-based X-noiseifier
LASSO	least absolute shrinkage and selection operator
MRI	Magnetic Resonance Imaging
RR	ridge regression
RSFC	resting-state functional connectivity
SC	structural connectivity
SVR	Support Vector Regression

Contents

1 Introduction

Human brain organization and behavior are inherently and intricately related. On one hand, behavioral traits like cognitive performance are often attributed to brain organization. Higher scores in cognitive tests may reflect more efficient information processing due to better organized brain networks. On the other hand, behavioral decisions and outcomes like alcohol consumption or sleep quality can influence brain organization in the long run. This complex relationship can also be observed between brain organization and clinical diagnosis. One fundamental aspect of neuroscience is to study these brain-behavior relationships, ultimately aiming to disentangle the underlying mechanisms or to inform clinical decisions.

Naïvely, brain-behavior relationships can involve assigning behavioral functions to brain regions. For instance, the motor cortex can be neatly divided into sections responsible for facial, arm, body and leg movements, each of which can be further divided into more fine-grained regions, such as the mouth, nose, and eyes for the facial section. However, when considering more complex cognitive or associative functions, such rigid assignments no longer suffice. Broadly speaking, we can relate the frontal cortex to executive function and working memory, the temporal cortex to memory, language and recognition, and the parietal cortex to attention modulation. Neither are these general relationships sufficient for uncovering the mechanism of these behavioral functions, nor are they specific enough for any clinical or practical application.

Complex functions intuitively require complex mechanisms. In order to understand complex cognitive or associative functions, neuroscientists started to consider the whole brain, or distributed brain networks, to be responsible for these functions. With each brain region as a node in a network, the brain can be seen as an information processing system, the efficiency of which can be related to a person's cognitive abilities. Following, the interactions between brain regions may be as important as, if not more important than, the individual functions of brain regions (Bressler, 1995, Tononi et al., 1998). The two perspectives represented by the highly specialized motor regions and the distributed

cognition network are not necessarily contradictory. They demonstrate the two principles of human brain organization: segregation and integration (Tononi et al., 1994, Friston, 2011, Fox and Friston, 2012), the combination of which forms the basis of our current understanding of the complex mechanisms of the human brain.

The interactions between brain regions, or brain connectivity, have been formally studied with different neuroimaging modalities and methods. These are mostly grouped under structural connectivity (SC), functional connectivity (FC), and effective connectivity (Friston, 1994). SC refers to the physical connections between brain regions, while FC refers to the statistical dependence between the signals observed from the different brain regions. Finally, effective connectivity measures aim to infer the causal effects between brain regions based on FC measures. The study of SC based on ex vivo neuroimaging data has been challenging, as the connections of interest lie beneath the cortex, i.e., inside the white matter. The study of FC, on the other hand, benefits from the more developed functional Magnetic Resonance Imaging (fMRI) techniques.

In particular, resting-state fMRI scans, where subjects lie in the scanner without performing any task (Biswal et al., 1995), are easy to collect from a large group of subjects (Arbabshirani et al., 2013, Scheinost et al., 2014, Finn, 2021). The resting-state functional connectivity (RSFC) is often thought to reflect not the brain connectivity specific to any task, but the intrinsic connectivity of the brain (Fox and Raichle, 2007, Mueller et al., 2013). It has therefore become popular to use RSFC for biomarker identification and psychometric prediction, assuming the existence of some connectivity-behavior relationships. In recent years, as more large samples of resting-state data became available, connectivity-based psychometric prediction (CBPP) has grown into its own field (Finn et al., 2015, Rosenberg et al., 2016, Dubois et al., 2018a, Dadi et al., 2019, Pervaiz et al., 2020, Yeung et al., 2022).

1.1 Functional Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a non-invasive scanning technique, based on the interaction of nuclei in atoms in human organs with the electromagnetic field in the scanner (Ombao, 2016). A nucleus with an odd number of protons and neutrons has a net nuclear spin, and will align its spin axis in the presence of an external magnetic field. For human brain imaging, the interaction is measured between the hydrogen atoms in

water and fats and the in-scanner electromagnetic field. Structural brain MRI images measured this way are useful for their high spatial resolution, as well as good distinction between different tissue types which contain different proportion of water and fats.

Following, fMRI was developed to provide proxy measures of functional activities in the brain, the most popular technique being the blood-oxygen-level dependent (BOLD) imaging. Here, we will only consider the BOLD technique, and use the terms fMRI and BOLD interchangeably. As the name suggests, the BOLD signals track the regional blood oxygenation level in the brain. It can be reasonably assumed that increased blood flow to brain regions reflects an increase in neuronal activity in the brain regions. The increase in blood flow leads to an increase in the ratio of oxygenated blood to deoxygenated blood in the region. Oxygenated hemoglobin take a longer time to lose magnetization and hence cause stronger BOLD signals, while deoxygenated hemoglobin result in weaker BOLD signals. Therefore, a stronger BOLD signal reflects an increase in blood flow, which in turn reflects an increase in neuronal activity in the brain region.

A popular fMRI design is task-based fMRI, where a behavioral task is performed by the participant while lying in the scanner. The neuronal responses represented by the BOLD signals during the investigated task can be compared to the signals during a baseline task. This contrast is considered the activation map for the investigated task. For example, to find the activation for working memory task, the contrast can be computed between the 2-back task signals and 0-back task signals. The 2-back task tests the participant's working memory by presenting a sequence of stimuli and, for each stimulus, asking the participant if it is the same as the stimulus from two stimulus ago, requiring both sustained attention and working memory. On the other hand, the 0-back task simply requires the participant to respond when a specific stimulus appears, involving sustained attention but not working memory (Miller et al., 2009). The brain regions activated for the 2-back task but not the 0-back task can hence be mapped to the working memory function specifically. It is common to establish mapping between brain regions and behavioral functions through activation studies based on task fMRI.

Alternatively, a more recent form of fMRI design called the resting-state fMRI has become widely adopted. As the participants only need to lie in the scanner, without falling asleep and without any task design, resting-state fMRI data are very easy to collect. The state of "rest" may be considered to be more similar to the everyday state of our brains. While we may not be actively performing any specific task, our brains are still supporting

our consciousness and are ready to start working on a task when the need comes. It is possible that human brains in this "resting state" show the intrinsic brain organization from which brain networks specific to tasks arise. As a result, resting-state fMRI data, especially RSFC, have been used to investigate a wide range of behavioral functions, as well as to probe the properties of the brain's intrinsic organization.

Compared to other popular techniques for studying brain functional activities like electroencephalography (EEG), BOLD images are lacking in temporal resolution and temporal accuracy. This is because blood oxygen level changes always are lagging behind actual activity onsets and offsets, doing so in a slow and smooth fashion. Commonly, a hemodynamic response function is used to model the changes in blood flow after the initial activity onset, describing the process of blood flow gradually increasing, peaking for a short duration, and then returning to baseline (Soares et al., 2016). It must be reiterated that fMRI techniques measures functional activities indirectly, in this case through the proxy measure of blood oxygenation level. Therefore, the regional BOLD signals are no reflection of how the brain regions actually function. The co-activation of a brain region based on fMRI data does not imply that the brain region is responsible for carrying out the mechanism of the investigated behavioral function. Direct brain-behavior relationships require more substantial evidence than only fMRI results to establish.

The advantage of fMRI is the finer spatial resolution, making it easier to localize the functional activities. While there are multiple levels and varying interpretation of functional units of the brain, some reasonable functional units to examine include neurons, cortical columns, brain circuits, and in some cases brain networks. The units (i.e., voxels) in fMRI images are, quite obviously, arbitrary. Depending on the exact spatial resolution, one voxel may correspond to multiple cortical columns and may hence cover multiple functional units. In addition, blood vessels do not supply blood to each cortical column separately, meaning that the actual activation in a neuron may lead to related blood flow changes in multiple voxels. This intrinsic smoothness of the fMRI data is widely recognized, and provides the basis of some fMRI data analysis techniques. For instance, the high similarity of neighbouring voxels lets us group them into parcels, so as to reduce the dimensionality of the data. Being able to compute parcel-to-parcel (instead of voxel-to-voxel) RSFC makes RSFC-based analysis and predictions much more viable to implement.

1.2 Connectivity-based Psychometric Prediction

CBPP studies explore brain-behavior relationships by relating inter-individual variability in RSFC to inter-individual variability in psychometric measures. Essentially, machine learning algorithms have been employed to predict behavioral performance, for instance, fluid intelligence (Finn et al., 2015, Ferguson et al., 2017, Greene et al., 2018, Gao et al., 2019, Maglanoc et al., 2020, Pan et al., 2021, Sen and Parhi, 2021). The standard protocol consists of first denoising the resting-state fMRI data, then parcellating the brain into brain regions with an atlas, followed by computing the Pearson's correlation between each pair of brain regions' mean timeseries. Finally, machine learning models are estimated using these correlation values as independent variables and the psychometric variables as dependent variables.

The prediction approach of investigating brain-behavior relationships is sometimes seen as an opponent to the association approach. In an association study, traditionally, a general linear model is fitted on the neuroimaging features and the behavioral measure(s), determining the features that are associated with the behavioral measure(s). In contrast, prediction models always need to be evaluated on unseen test data not used in estimating the model, either using a cross-validation scheme or a external set of validation data. This makes the prediction approach more exciting for its potential utility, especially for clinical applications (Bzdok et al., 2021). However, it has become more and more common to use machine learning algorithms with cross-validation or external validation in association studies to find more generalizable associations (Genon et al., 2022). Fundamentally, it should not be forgotten that both approaches examine the relationship between inter-individual variability in neuroimaging features and that in psychometric variables. Despite methodological differences, insights discovered by the two approaches should validate each other rather than contradicting.

Not unlike other fields, CBPP studies were pioneered by promising yet over-optimistic results. Conceptually, CBPP is promising for both practical applications and revealing neurobiological insights (Haufe et al., 2014, Bzdok and Ioannidis, 2019). Nevertheless, the assumption of the existence of connectivity-behavior relationships is easily forgotten, or often ill-defined. The intricate nature of brain-behavior relationships implies that the relationship between any connectivity feature and psychometric variable may be mediated by confounding factors, such as age, or education. Demographic and

physiological factors can influence both RSFC and the psychometric variable, making it hard to extract neurobiological insights from the relationships learnt by the model. Furthermore, if the contribution of neurobiological factors cannot be assessed independently from confounding factors, the usefulness of CBPP cannot be properly established. Practically, an individual's cognitive abilities can be estimated based on the individual's education level or cognitive impairment based on age. The utility of any CBPP model has to be based on additional predictive power not already explained by non-neurobiological features.

Moreover, each step of the standard procedure may be implemented differently in different studies. Due to the large number of options available at each step, there is no guarantee that prediction outcomes from different studies are comparable. This is further complicated with the use of different datasets involving cohorts of varying demographics, as well as data collected and processed by differing teams. These concerns motivate the call for model interpretability, requiring the model(s) to be examined not only for the prediction accuracy, but also the brain-behavior association derived. Potentially, such models may allow the contributions of confounding factors, implementation decisions, cohort idiosyncrasies, and the neurobiological features to be disentangled. Most crucially, the impact of these factors should be assessed and represented transparently.

This work attempts to disentangle some of the factors involved in, or confounding, brain-behavior relationships, using a region-wise approach. For each brain region, a machine learning model is estimated and tested for psychometric prediction using the region's RSFC to all other regions. The relative predictive accuracies of the region-wise models are hence a direct reflection of their connectivity profile's involvement in predicting the target psychometric variable. The distribution of these model accuracies across the brain was observed to change when certain confounding factors were controlled, or when different denoising strategies were used, capturing the effect of these factors. Furthermore, using this distribution to represent the brain-behavior association derived from the models, we then demonstrated that achieving similar prediction accuracies in different cohorts does not mean the same brain-behavior associations were learnt. While some consistent patterns of brain-behavior association can be identified using different cohorts, the consistency relies on the similarity of data collection and processing protocols used in the different cohorts. Overall, we demonstrated how interpretable models can help to disentangle factors confounding brain-behavior

relationships learnt, which in turn improves the neurobiological insights that can be extracted from the models.

1.3 Factors Influencing Brain-behavior Relationships Modelled

1.3.1 Denoising

The fMRI technique measures brain activity indirectly using the BOLD signals, which are affected by both neural and non-neural sources (Power et al., 2017). Therefore, before processing any fMRI data, preprocessing is needed to remove the noise from non-neural sources. These sources commonly include head motion, imaging artifacts, as well as signal location (Esteban et al., 2019).

The first step of preprocessing usually includes motion correction, slice-timing correction, and susceptibility distortion correction. It may also be recommended to remove the first few frames (or time points) in the fMRI data, if they are detected as non-steady state volumes. The sequence of these steps can be different in different preprocessing pipelines and implementations. In particular, there is no consensus on whether motion correction or slice-timing correction should be performed first. As for susceptibility distortion correction, it is not always possible to perform this correction as it requires field maps or pairs of scans with opposite phase encoding directions (Poldrack et al., 2011, Glasser et al., 2013). Without susceptibility distortion correction, signal loss or distortion may be commonly observed in the anterior prefrontal cortex and orbitofrontal cortex (Poldrack et al., 2011).

Next, the fMRI image need to be aligned to the anatomical T1w image, and to a standard template. The co-registration to the anatomical image is usually done with simple linear registration as the native fMRI image and the T1w image of the same individual are expected to be similar. The T1w image can be aligned to the standard space with an affine registration followed by a nonlinear registration (Jenkinson et al., 2002, Andersson et al., 2007, Avants et al., 2009). Alternatively, the fMRI data can be mapped onto a surface mesh, which represents the cerebral cortex as a 2-dimensional sheet. The surface system may be preferred for respecting the cortical topography (Anticevic et al., 2008, Ghosh et al., 2010, Tucholka et al., 2012), but suffers from not being applicable to subcortical regions in many cases. The accuracy of spatial alignment during preprocessing directly affects the quality of the neuroimaging features extracted

based on the transformed images.

The fMRI data in the standard space can be considered minimally usable, although further denoising would be done in most cases. For instance, motion scrubbing may be done by regressing out volumes with high motion, simply removing them, or re-estimating these volumes based on their neighbors (Lemieux et al., 2007, Power et al., 2012, Satterthwaite et al., 2013). Another group of popular techniques is the Independent Component Analysis (ICA) based denoising strategies, namely the FMRIB's ICA-based X-noiseifier (ICA-FIX) and ICA-based Automatic Removal of Motion Artifacts (ICA-AROMA). These techniques involve extracting ICA components from the fMRI data and then removing the noisy components, which may be motion-related noise or artifacts from other sources, based on a previously trained machine learning model (Griffanti et al., 2014, Pruim et al., 2015b, Pruim et al., 2015a). As the last step of preprocessing, a standard practice is to implement nuisance regression to control for signals from non-neural nuisance sources, including the 24 motion parameter, white matter signals, cerebrospinal fluid signals, and occasionally the global signal.

Regarding the more sophisticated denoising techniques, there is no consensus or evidence on which one should be preferred. Some may argue that motion scrubbing introduces artificial modifications to the fMRI timeseries, potentially disrupting the temporal information in the data. The reliance of the ICA-based techniques on previously trained models means that the noise identified may be affected by the type of noise existing in the data used in these previously trained models. Finally, while the neurobiological implications of removing the global signal are still not clear, it has been shown that global signal regression may improve CBPP performance for many psychometric variables (Li et al., 2019).

On one hand, it is intuitive that difference in preprocessing choices would impact the brain-behavior relationships learnt by the model. On the other hand, the influences of preprocessing choices on the neuroimaging features are generally not straightforward. More research is still required to elucidate the effect of each preprocessing step on the neurobiological information in the fMRI data processed.

1.3.2 Brain Atlases and Connectivity Computation

Before connectivity can be computed between brain regions from the fMRI data, the location and boundary for each brain region of interest must be defined. As it is not computationally feasible to use voxel-wise (or vertex-wise) data for prediction purposes, the data dimensionality needs to be reduced by defining brain regions encompassing groups of voxels (or vertices). This is typically done using a group atlas, providing the same definition of brain regions (or parcels) for all subjects in a study. Alternatively, individualized parcellations can be used to define brain regions suiting each subject's topography separately. The latter can cause heavy computation and storage demands, but have been shown to improve CBPP accuracies slightly (Kong et al., 2019, Kong et al., 2021).

The choice of the group atlas to use could impact both the prediction performance and model interpretation in a CBPP study. First, the coordinate system where the atlas was developed and the brain coverage need to be appropriate for the data present. While many researchers prefer the surface coordinate system, surface-based data and atlas usually suffer from a lack of subcortical coverage, as well as surface mapping failures when data quality is sub-optimal. Second, the granularity (i.e., the number of parcels) of the atlas to use reflects the assumption made on the number of functional brain regions in the human brain. In particular, for the cerebral cortex, the estimate for the number of brain regions is between 300 to 400 (Van Essen et al., 2012). Finally, the modality of the data used for estimating the atlas should, intuitively, match the modality of the data used for prediction.

With a chosen atlas, one summary signal needs to be extracted for each parcel. The simple and common way is to compute the average BOLD signal across all voxels or vertices inside a parcel. Alternatively, if an ICA map is used for parcellation, where components can overlap with each other, dual regression is usually employed for estimating the summary signal (Hacker et al., 2013, Nickerson et al., 2017). In either case, the data dimensionality is reduced to the number of parcels $\times$ the number of time points. Then, FC can be computed between all pairs of parcels, resulting in a connectivity matrix with the size of the number of parcels $\times$ the number of parcels. In a typical whole-brain CBPP model, the input features include the upper or lower triangular part of the connectivity matrix, excluding the diagonal elements.

FC is most commonly computed as the Pearson's correlation coefficient between two brain regions' fMRI signals, i.e.,

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}},$$

where r is the correlation coefficient, and x and y are the two signals. To make sure that FC values are comparable across different subjects, the Fisher's z transformation is applied to each correlation value. This transforms each r value into

$$z = \frac{1}{2} \ln(\frac{1+r}{1-r}),$$

leading to a approximately normal distribution of the transformed correlation values. Another technique for computing connectivity is partial correlation, where each correlation coefficient between a pair of brain regions is computed while controlling for the effects of all other brain regions' signals. This method may be considered as measuring only the direct connection between two brain regions, removing the effect of indirect connection through other brain regions (Bijsterbosch and Beckmann, 2017, Lim et al., 2019). Furthermore, the connectivity matrix can be projected into tangent space, which allows the application of classification algorithms requiring a Riemannian manifold, as well as improving the prediction power (Pervaiz et al., 2020).

The neurobiological implications of the tangent space projection, and to some extent the partial correlation method, are not well understood. While both have been shown to improve prediction performance, they may be less desirable when interpretability is concerned. For instance, the individual connectivity values in tangent space may not represent the connection between the pairs of brain regions anymore.

1.3.3 Machine Learning Models

CBPP studies generally make use of regression models to predict continuous psychometric variables. The type of regression algorithm used include simple linear algorithms like multiple linear regression and ridge regression (RR), sparse algorithms like least absolute shrinkage and selection operator (LASSO) and relevance vector regression, more complex linear algorithms like elastic net (EN), as well as nonlinear algorithms like kernel regression and deep neural networks.

When a machine learning model is trained, it learns the parameters defining a relationship between the input features and the target psychometric variable to predict. In the case of a linear model, the relationship is simply,

$$\hat{y_i} = \sum_p w_p x_{p,i} + c,$$

where y_is are the target psychometric variable for subject i, $x_{p,i}$s are the pth input features for subject i and w_ps the corresponding weight, and c is a constant. A more complex model may consider nonlinear terms, e.g.,

$$\hat{y_i} = \sum_p (w_p x_{p,i} + u_p x_{p,i}^2) + \sum_{p,j \neq i} v_p x_{p,i} x_{p,j} + c,$$

where u_ps and v_ps are the weights for the nonlinear terms. In the simplest case, the weights in the equation are estimated by minimizing the squared error, with a loss function of

$$L = \sum_i (y_i - \hat{y_i}).$$

Additional terms can be added to the loss function to enforce specific constraints and to apply regularization to reduce overfitting.

The choice of the model to use in a CBPP study hence reflects the assumption made on the relationships between the neuroimaging features and the psychometric variable. As it is common to assume a linear relationship between neuroimaging features and psychometric variable, linear models are usually considered more biologically plausible. Nevertheless, nonlinear models may be preferable when prediction performance is the main goal. Ideally, a nonlinear model should be able to capture the actual brain-behavior relationships with sufficient training data in a data-driven manner, even when the true underlying relationship is linear.

Moreover, the assumption made on the brain-behavior relationships is reflected by the confound controlling strategy. A confound is a variable that influences both the independent and the dependent variables. Not accounting for confounding effects will hence lead to false positives in the brain-behavior relationships derived, which are actually confound-behavior relationships. Common confounding variables considered in CBPP studies include age, gender, handedness, brain size, and in-scanner movement. Usually, the confounding variables are controlled by regressing them out from the input features,

harmonization, stratified sampling, or re-weighting. Sometimes, when the confounding variables also contain some neurobiological effects of interest, it may also be desirable to include the confounding variables as input features. Overall, the way that confounds are dealt with is essentially a reflection of the assumed causal relationships between the input neuroimaging features, the confounding variables, and the psychometric variables.

Another contention in model selection is whether to impose sparsity restrictions or not. The concept of sparsity involves picking only a small subset of input features to be actually used in model, assuming that the actual relationship between the input features and the target psychometric variable only involves this subset of input features. The validity of this assumption is naturally dependent on the type of neuroimaging data and psychometric variable in question. Again, a specific motor function may be related to certain features from a specific brain region, while higher cognitive functions tend to involve a distributed network of brain regions. In the latter case, it may not be suitable to use a model that enforces sparsity. A popular algorithm in CBPP studies, EN, makes a compromise between the L1 (or ridge) regularization constraints, which helps to prevent overfitting, and the L2 (or lasso) regularization constraints, which promotes sparsity, resulting in a loss function in the form of

$$L = \sum_i (y_i - \hat{y}_i)^2 + \alpha\lambda \sum_p |w_p| + \alpha(1-\lambda) \sum_p w_p^2,$$

where α determines the extent of regularization in general, and λ determines the trade-off between L1 (second term) and L2 (third term) regularization. The extent of sparsity is determined by tuning the hyperparameter λ using nested cross-validation in the training set, making no prior assumption on sparsity.

Finally, overfitting is a general issue in machine learning, when the model has captured so much information specific only to the training data that it fails to generalize to the test data. This can happen due to many reasons, for example, the lack of regularization, an overly high number of model parameters, and an insufficient amount of training data. When dealing with RSFC features, the input feature dimension is usually in the order of thousands to tens of thousands. Assuming no sparsity constraint, at least one model parameter needs to be estimated for each input feature, making the number of model parameters in RSFC-based predictions inherently high. To prevent overfitting, either the model implementation needs to be specially designed for

high-dimensional data, or feature selection needs to be carried out before training the model to reduce the number of input features.

A popular method of feature selection is to compute the Pearson's correlation between the input features and the target psychometric variable in the training data, selecting features where the correlation is statistically significant or selecting the top features ranked by the correlation coefficients (Finn et al., 2015, Dubois et al., 2018a, Yeung et al., 2022). Similar to the sparsity constraints, this approach relies on the assumption that complex brain-behavior relationships are not underrepresented by the reduced set of features. In our work, we approach the overfitting issue from a different angle. By estimating a model for each brain region using only the connectivity between the region and other brain regions, the number of features for each region-wise model is only in the order of hundreds. This helps to prevent overfitting without the need for feature selection. Finally, brain regions where the region-wise model performed relatively better than the others can be picked. Features from these brain regions can be pooled for estimating a final model, accounting for interaction of features between brain regions and hence optimizing for overall prediction performance.

Machine learning often seems to promise data-driven solutions. In ideal scenarios, machine learning models should be able to learn the underlying relationships automatically with enough training data, computational power, and time. During actual implementations concerning neuroimaging data, however, human decisions are usually still required to train a proper model. As every decision made when building the model can affect the assumed brain-behavior relationship to learn, it is crucial to record the model parameters faithfully and present them transparently. Furthermore, in order to understand what brain-behavior relationships are captured by the model, interpretable models may be preferable sometimes, in comparison to the powerful but less interpretable deep neural network models. It may be tempting to follow the hype of deep learning which has achieved substantial success in the field of computer vision as well as natural language processing. When studying brain-behavior relationships, nevertheless, black-box models which are difficult to interpret and to trust can be dangerous in actual application. Unfair models may wrongly associate intelligence with racial or gender factors, while clinical models plagued by confounds may cause wrong treatments to be delivered to patients. In the long run, the CBPP field must proceed cautiously, to eventually bring predictive modeling to practical usefulness.

1.4 Issues and Potential Solutions

Despite the potential of CBPP approaches in terms of practical utility and providing neurobiological insights, they have been limited in several aspects in the current state. First, the prediction accuracies tend to be low when sufficiently large datasets are involved (Sui et al., 2020, Genon et al., 2022, Yeung et al., 2022), raising questions about the applicability of CBPP models to the general population. Many studies have investigated factors and potential solutions for improving psychometric prediction performance (Dadi et al., 2019, Li et al., 2019, Pervaiz et al., 2020, Finn and Bandettini, 2021, Rasero et al., 2021), which will not be the focus in this work. Second, CBPP models tend to be difficult to interpret due to their multivariate nature, limiting both the validity of the models themselves and any potential insights to be derived from the models (Scheinost et al., 2019). Third, the replicability of CBPP results and the generalizability of CBPP models across different neuroscience cohorts have not been established (Sui et al., 2020, Genon et al., 2022).

1.4.1 Interpretability

The difficulty in interpreting CBPP models is a result of the difficulty in quantifying each feature's contribution to the prediction, as well as the difficulty in the meaningful representation of RSFC features in alignment to brain mapping literature.

A popular yet misleading approach to quantify each feature's contributions is to simply use the weight assigned to each feature by the machine learning algorithm, where a larger weight signifies higher importance of the feature. Noting that almost all machine learning models employed in CBPP approaches were discriminative, or backward, models, the relationship between the weight and feature importance may actually be the opposite (Haufe et al., 2014). An alternative would be to use the weights assigned by a sparse algorithm to represent binary feature contribution; a sparse algorithm only assigns weights to a small set of features, which can then be considered contributing features. Nevertheless, such interpretations are inherently dependent on the designs of the machine learning algorithms. As RSFC features are highly correlated, an outcome of the complexity in brain-behavior relationship itself (Genon et al., 2018a, Genon et al., 2018b), they may cause weight assignment to be unstable (Zou and Hastie, 2005, Hastie et al., 2009).

The region-wise CBPP framework proposed in this work is designed to directly quantify each brain region's contribution to the prediction. A region-wise model is estimated and evaluated using each brain region's RSFC profile (i.e., connectivity edges between this region and all other regions), while a whole-brain model is estimated and evaluated using all region-to-region RSFC. Each brain region's contribution to the prediction is, therefore, the ratio of its region-wise model's prediction accuracy to the whole-brain model's accuracy. Besides the straightforward and data-driven interpretation, the region-wise approach is also intended for easier and meaningful representation of the findings. By aggregating edge-based features by brain regions, visualization of region-wise prediction contribution becomes really convenient. The set of relevant brain regions, where relatively higher region-wise contribution was observed, can be compared with brain mapping literature for validation.

1.4.2 Replicability and Generalizability

External validation is an important concept in predictive modeling, as a model can only be predictive (and useful for practical applications) if it can predict outcomes in unseen data (Scheinost et al., 2019, Poldrack et al., 2020). It is not necessary to implement such validations in completely new samples, as cross-validation within a single cohort is possible and usually the more feasible choice (Yeung et al., 2022). Nevertheless, As many large-scale neuroscience datasets has become publicly available (Nooner et al., 2012, Van Essen et al., 2013, Sudlow et al., 2015, Karcher and Barch, 2021), it is now viable to assess the cross-cohort replicability and generalizability of CBPP models.

Replicability and generalizability are two aspects of the same issue. Replicability means that the results obtained from CBPP models in one cohort should be similar to those obtained in a different cohort. From the perspective of the neuroscience field, this also means that results obtained from one study should be replicable by another research group. Generalizability means that the CBPP models estimated based on one cohort should be able to predict outcome in a different cohort too. This is often achieved by evaluating the model in out-of-sample test data (Beaty et al., 2018, He et al., 2020, Jiang et al., 2020).

Furthermore, replicability and generalizability of CBPP models can be investigated from two directions, focusing on either the prediction accuracy or the derived brain-behavior association pattern. In this work, the focus is on the less

examined latter direction. As the region-wise prediction contribution map can be used as a representation of the brain-behavior association pattern derived from the region-wise CBPP framework, the replicability and generalizability of the patterns can be simply computed as correlation between the relevant maps. For instance, replicability of the prediction pattern may be evaluated as the correlation between prediction maps generated in different cohorts. Generalizability may be evaluated by estimating a prediction model in one cohort and test it in another, where generalizability can thus be measured by the correlation between the prediction map generated in the training cohort and that generated in the test cohort. The motivation is that the replicability and generalizability of interpretations of CBPP models are essential to their utility in potential practical applications. Addressing these issues are not only beneficial to the validity of the region-wise CBPP framework, but also to the field of psychometric prediction in general.

1.5 Ethics Protocols

The ethics protocols were approved by the Ethics Committee of Heinrich Heine University Düsseldorf (4039 and 2018-317-RetroDEuA).

1.6 Aims of Thesis

The aim of this work is to improve the interpretability of connectivity-based models for psychometric prediction. The interpretability here refers to the ability to reveal brain-behavior associations based on psychometric prediction results. Conceptually, this is achieved by implementing prediction for brain regions separately, in the proposed region-wise CBPP framework. The framework hence allows interpretation based on each brain region's relation to the investigated behavior. The research question in interest is the attribution of behavioral traits to sets of brain regions, identified in a data-driven manner.

In study 1, the framework was established and demonstrated with two applications, presenting brain-behavior relationships by constructing a psychometric prediction profile for each brain region, as well as by constructing a prediction performance variation map across brain regions for each psychometric variable. To validate the framework, the brain-behavior associations derived are aligned with brain mapping literature. In study 2, the framework was utilized so as to investigate the replicability and generalizability of CBPP, in terms of both prediction performance and the brain-behavior relationships derived,

in multiple cohorts. Overall, the region-wise CBPP framework addresses the concerns of CBPP model interpretation, demonstrating the validity and replicability of certain brain-behavior relationships derived from the framework. In study 3, we performed a literature survey on existing psychometric prediction studies based on neuroimaging data. By reporting the trends across these studies, known and new issues can be identified and addressed in future studies.

2 A Connectivity-Based Psychometric Prediction Framework for Brain-Behavior Relationship Studies, Wu, J., Eickhoff, S.B., Hoffstaedter, F., Patil, K.R., Schwender, H., Yeo, B.T.T., Genon, S., Cerebral Cortex, 31: 3732-3751, (2021)

3 Cross-Cohort Replicability and Generalizability of Connectivity-Based Psychometric Prediction Patterns, Wu, J., Li, J., Eickhoff, S.B., Hoffstaedter, F., Hanke, M., Yeo, B.T.T., Genon, S., NeuroImage, 119569, (2022)

4 Reporting Details of Neuroimaging Studies on Individual Traits Predictions: A Literature Survey, Yeung, A.W.K., More, S., Wu, J., Eickhoff, S.B., NeuroImage, 119275, (2022)

5 Discussion

In this work, we have developed and validated the region-wise CBPP framework focusing on interpretability. In study 1, we developed the framework based on a preliminary evaluation of existing approaches and parameters used in previous CBPP studies. We then demonstrated two applications of the region-wise framework, for constructing individual brain region's psychometric profile and for constructing individual psychometric variable's prediction pattern. Both psychometric profiles and prediction patterns can be aligned to brain mapping literature. Furthermore, they were helpful in investigating the effect of denoising strategies and confound controlling strategies. In study 2, we investigated the cross-cohort replicability and generalizability of the prediction pattern of fluid cognition and openness. We demonstrated the usefulness of the region-wise framework in assessing the prediction replicability and generalizability in terms of brain-behavior association derived from the prediction models, and not only in terms of prediction accuracy. While a low to moderate extent of replicability and generalizability was found for fluid cognition predictions, we noticed that higher replicability and generalizability inevitably rely on higher similarity in terms of data collection protocols in the different cohorts. Finally, in study 3, we found a negative correlation between prediction accuracies and the training sample size, as well as a lack of proper cross-validation implementation and external validations. These trends may be further reflections on the prevalence of overfitting problems and a general lack of generalizability in existing CBPP models developed.

5.1 Region-wise Psychometric Profiles and Psychometric-wise Prediction Patterns

The region-wise psychometric profiles are constructed by evaluating region-wise models across a range of psychometric variables. A brain region's associations with different types of behavioral measures are hence represented by its region-wise model's varying predictive power for these behavioral measures, presented in study 1 in a radar plot. We demonstrated that while similar psychometric profiles can be found between parcels in the same location but different hemispheres, the small differences in their psychometric profile

can be related to their functional differences revealed in brain mapping literature. For instance, while both the left and right Broca region showed predictive power for working memory, the left Broca showed higher predictive power than the right Broca for language task performance. This is in line with brain mapping literature, showing the involvement of both Broca regions in phonological memory (Salmon et al., 1996, Smith and Jonides, 1999, Zurowski et al., 2002, Rogalsky and Hickok, 2011), as well as the left Broca region's specific involvement in language function (Clos et al., 2013, Karolis et al., 2019).

The psychometric-wise prediction patterns are constructed by evaluating region-wise models from every brain region for a single psychometric variable. A psychometric variable's associations with different brain regions are hence represented by the predictive power of these different brain regions, presented in study 1 as a map of prediction accuracy distribution across the brain. Again, we demonstrated that these prediction patterns are well aligned to brain mapping literature. For instance, while the prediction pattern for 2-back task performance is similar to the prediction pattern for 2-back task performance only involving face tasks, the latter revealed additional predictive power in the visual regions, as well as right ventral temporal regions responsible for face processing (Sams et al., 1997, Nakamura et al., 2000. Nelson, 2001).

It is important to note that full alignment should not be expected between prediction patterns (or psychometric profiles) and brain mapping literature. This is because brain mapping results are usually based on activation experiments, focusing on regions that consistently activate for a behavioral test. On the other hand, prediction models relate individual variability in neuroimaging features to individual variability in behavioral measures, and hence tend to highlight features with high individual variability. Conceptually, the brain regions that are identified by both the region-wise approach and brain mapping literature may be consistently involved in the behavioral function, but in different ways or to different extents in different participants. The brain regions identified in the prediction pattern that do not align with brain mapping literature may still be related to the predicted behavior, but not consistently so in every individual. The brain mapping literature is not necessarily the ground truth to compare the prediction patterns against. Ultimately, the validity of the brain regions identified by the region-wise approach need to be assessed based on known neural mechanisms, as well as fully explainable model mechanisms.

5.2 Region-wise and Whole-brain Approaches in Complement

The region-wise CBPP framework was developed not to be a competitor to whole-brain approaches, but to be used in complement with the latter. Considering complex behavioral functions arising from distributed brain networks, a region-wise model estimated using only connectivity to one brain region would not have been optimal for prediction performance. Instead, the best way to model such behavioral functions would be to include all potentially useful features in a global (or whole-brain) model, as well as modeling the interaction between these features. In this case, the region-wise approach can still be useful in two ways. First, the region-wise models can be estimated and validated on the training set. Regions where the region-wise models obtained statistically significant prediction accuracies can be selected to use in the global model. Second, for the selected regions, their region-wise model can be evaluated in the test set along with the global model. The prediction accuracies of these region-wise model can be compared to the global model, resulting in a fractional value of the ratio of region-wise model accuracy to global model accuracy.

The region-to-global fractional value represent the contribution of the RSFC features of a brain region, in the context of the global model. As RSFC features are highly correlated, it is possible for multiple brain regions to have similar contribution to the same psychometric variable. In our examination of complex behavioral measure, represented by the cognition composite scores, we found widespread regions with moderate contribution. Mostly, these regions showed accuracies that are about half of the whole-brain model accuracies. This aligns with the assumption that the composite measures of cognition are related to distributed brain networks, and not dominated by any individual brain region (Dubois et al., 2018b). On the other hand, for more specific measures like cognitive flexibility and working memory performance, we found that certain region-wise models could perform as well as the whole-brain models. This may suggest that interindividual variability of behavioral functions like working memory may depend on a specific network of connected brain regions.

In a follow-up analysis, we also found cases where certain region-wise models performed better than the corresponding whole-brain model (figure 1, figure 2). In particular, all best region-wise models across different granularities were better than the corresponding whole-brain model for several measures including agreeableness, emotion

recognition, anger affect, fear affect, episodic memory (picture sequence), executive function (card sort), processing speed, and dexterity. This may suggest that these psychometric measures are related only to a small subset of all the RSFC features. The region-wise models benefited from the reduced number of features, for both reduced overfitting and underfitting. Further investigation could analyze if whole-brain models for these psychometric measures can be improved by only including features from the significant region-wise models.

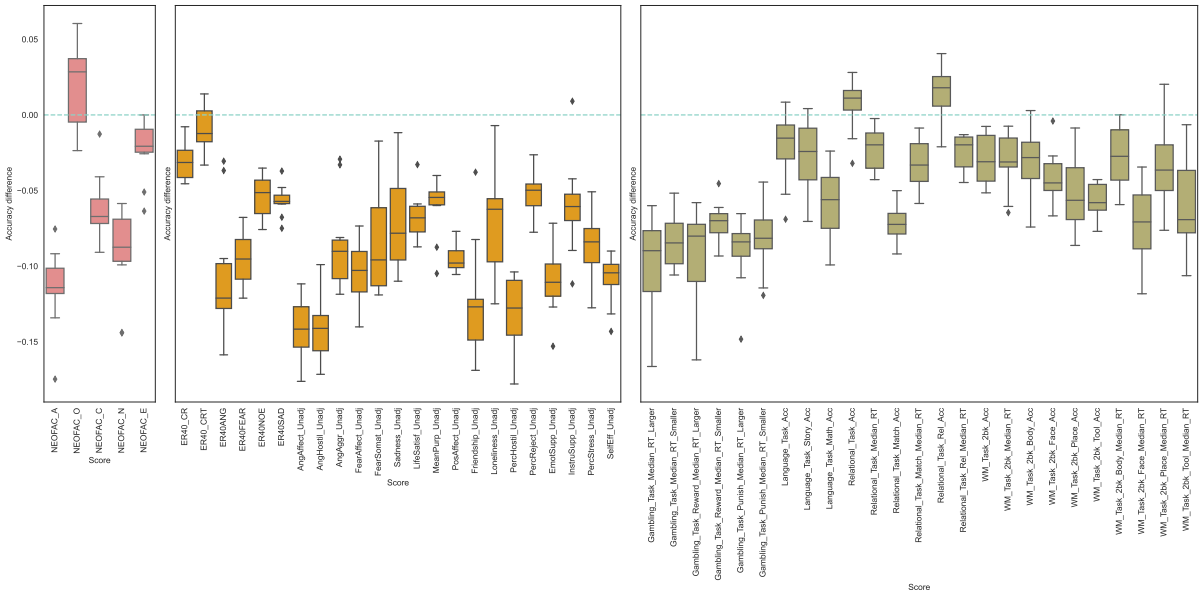

Figure 1: **Accuracy differences between the whole-brain model and the best region-wise models for personality (red), emotion (orange), and in-scanner task (olive) measures.** Black bars inside the boxes represent the median accuracy different across different parcellations. Black diamond-shape markers represent outliers.

5.3 Effects of Prediction Implementation Decisions

Our results provided evidence for the benefit of ICA-FIX for prediction studies, both in terms of prediction accuracy and in terms of brain-behavior associations derived. For the prediction patterns derived, models using ICA-FIX denoised data showed significantly higher variance across parcels than models using minimally processed data. For both psychometric profiles and prediction patterns, the Euclidean distances between profiles across psychometric variables or between patterns across parcels were also significantly larger using ICA-FIX denoised data. We hence conclude that ICA-FIX denoising helps to preserve interindividual variability in regional RSFC features while removing global artifacts, making it easier to establish psychometric profiles specific to individual parcels and prediction patterns specific to individual psychometric variables.

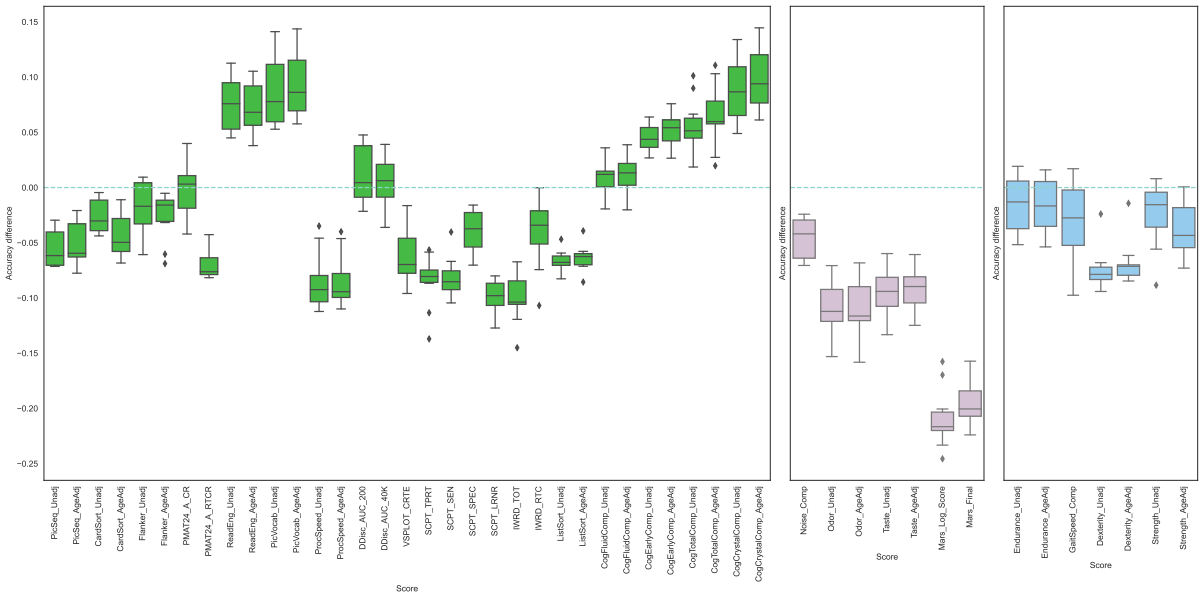

Figure 2: **Accuracy differences between the whole-brain model and the best region-wise models for cognition (green), sensory (purple), and motor (blue) measures.** Black bars inside the boxes represent the median accuracy different across different parcellations. Black diamond-shape markers represent outliers.

When examining surface-based parcellations of different granularity level (100-parcel to 400-parcel), we found that more parcels with relatively better predictive power can be identified at 200-parcel and 300-parcel granularity. For both replicability and generalizability analyses, better results were obtained with higher granularity (232-parcel, 350-parcel, and 450-parcel) than lower granularity (116-parcel). These evidence aligns with the commonly acknowledged hypothesis stating that there are 300-400 cortical brain regions (Van Essen et al., 2012). While these brain regions may still be divided into smaller functional units, our results suggest that a granularity of about 300 parcels would lead to the optimal representation of functional units for CBPP studies. It is also possible that higher granularities leads to better predictions in general. Future analyses could assess the effect of granularity on prediction performance up to, for instance, 1000 parcels. However, it should be noted that higher granularity also means larger connectivity matrix and hence higher number of features. Adding more features with higher granularity may not help to improve prediction performance, if feature selection is required to restrict the number of features.

Regarding RSFC computation methods, we found that partial correlation improved prediction performance if ICA-FIX denoising was not applied, but mostly results in similar prediction performance as Pearson's correlation otherwise. A plausible explanation is that the regression step in computing partial correlation removed similar

global artifacts that ICA-FIX did. Nevertheless, we also found that partial correlation causes regional psychometric profiles to be very similar to each other, possibly due to the aggressive regression. As RSFC features are high correlated, suggesting that parcels in similar functional networks may have similar connectivity profiles naturally, removing the common variance between these parcels during partial correlation may have caused loss of region-specific information. As a result, Pearson's correlation may still be the best choice for computing RSFC if interpretability is the concern.

Lastly, we tested different linear models and found that a simple implementation of the Support Vector Regression (SVR) is as good as the more complex EN, when prediction accuracies were assessed as the Pearson's correlation between predicted and observed values. However, if an accuracy measure assessing absolute error was used, EN and RR are then the best choice most of the time. The lower accuracies from SVR is expected as the implementation of SVR used the 'fitrlinear' function from Matlab, which is highly efficient with high-dimensional input features but may compromise the prediction accuracies. For RR and EN, a potential factor influencing brain-behavior relationships that was not investigated in our study would be the hyperparameter tuning method. As the hyperparameter tuning steps in both study 1 and 2 were done by optimizing the correlation-based accuracy, further investigations would be needed to know if the hyperparameter tuning implementation by absolute error-based accuracy would influence the prediction accuracies.

5.4 Region-wise Approaches for Feature Selection and Data-driven Discovery

In order to train a proper model in the presence of highly correlated RSFC features, the number of features should be restricted to roughly $\sqrt{N}$ in a dataset of N subjects (Jain and Waller, 1978). This limits the application of CBPP models to clinical population or locally acquired datasets where smaller sample sizes are more common, leading to higher risk of overfitting and less generalizable models. Usually, a priori knowledge is used to select regions of interest, based on canonical networks defined based on meta-analysis or RSFC-derived atlases (Nostro et al., 2018, Chen et al., 2021, Pläschke et al., 2020). However, our results demonstrated that brain regions that are functional hubs and/or with high interindividual variability in terms of RSFC tend to have predictive region-wise models, at least for cognition measures. These regions may be overlooked in canonical networks

as they are not directly related to the target psychometric variable. In comparison, we consider the region-wise CBPP framework a more systematic and data-driven approach for feature/region selection.

Our results of key regions identification highlighted the importance of hub regions in psychometric predictions, potentially suggesting the crucial role of information processing in cognitive tasks. The widespread set of brain regions identified for cognition composite scores, each contributing up to half of the accuracy of the whole-brain model, may support the global view of brain-behavior relationships where distributed networks and not localized regions are responsible for behaviors. While our results are not direct evidence for confirming any hypothesis, they could help to guide future investigations.

Nevertheless, it should be noted that feature selection based on region-wise CBPP is still limited by the generalizability of the region-wise models themselves. In our examination of fluid cognition prediction, we found a set of key brain regions with decent generalizability. On the other hand, in our examination of the openness measure, which is also one of the most well predicted psychometric variables in healthy population, we found little replicability in the region-wise prediction patterns across different cohorts. Therefore, the key regions selected by the region-wise framework in one study sample cannot always be assumed to be useful in a different sample. Potentially, key regions where the relationships with the psychometric variable can be theoretically explained may be safer to use for guiding feature selection in a new sample.

5.5 General Limitations of the Field

A major critique of the CBPP field as a whole is the low prediction accuracies. This poses limitations on both potential applications of the developed models as well as the validity of model interpretations. After all, the brain-behavior relationships learnt by a model can only be worth interpreting if the model can actually predict the target behavioral measure sufficiently. The low prediction accuracies are partially due to the limited amount of variance shared between RSFC and psychometric variables in healthy population (Scheinost et al., 2019), independent of demographic factors like age. For practical utility, it may be worthwhile to assess the prediction performance of multimodal prediction combining features from task, structural and diffusion MRI. Further along the line, complementing data using MRI with data using other imaging techniques (e.g., EEG, functional near-infrared spectroscopy, and positron emission tomography) may be

considered.

Other possible causes of the low prediction accuracies include low reliability of the psychometric measures, noisy neuroimaging data, and insufficient sample size. The quality of the psychometric and neuroimaging data relies on the data collection team, any preprocessing applied, as well as any inherent shortcomings of the behavioral tests or imaging techniques performed. While some are pessimistic about the future of psychometric prediction because of their lack of confidence in the psychometric and neuroimaging data, others suggest that predictions will still be viable with larger sample size, evidenced by the effect of big data in the field of computer vision. Nevertheless, big data inevitably mean big computational cost, placing a financial constraint on future research.

Finally, the limited replicability of prediction results and generalizability of prediction models to new cohorts remain to be tackled. Our results suggested that cross-cohort replicability and generalizability may depend on the similarity of the cohorts, in terms of data collection and processing protocols. By standardizing these protocols in the community, it may be possible to generally improve the replicability and generalizability. However, certain requirements on imaging resolution, motion artifacts controlling, and scan duration may be difficult to meet, especially in children, older adults, and clinical population. To this end, better preprocessing strategies that can standardize imaging quality post-hoc may be the potential solution.

5.6 Towards Multi-cohort Analysis and External Validation

In study 3, we found a curious negative correlation between prediction accuracy and sample size similar to Sui et al., 2020, mostly for studies using cross-validation in single cohorts. This is likely a result of overfitting models and peculiar subject selection in smaller samples, which also tend to be locally recruited samples. The higher prediction accuracies in smaller samples do not necessarily imply better trained models, but possibly reflections of publication bias. As only 25% of the surveyed studies used external test sets, we cannot assess the generalizability of the models with exceptionally high accuracies to test if they were indeed overfitting. Nevertheless, this negative correlation was not statistically significant between external validation accuracy and the external test sample size, suggesting that the use of external validation may be the way to address this issue.

In general, using large open datasets can be helpful in providing more generalizable

results, as well as external or hold-out test sets (Turner et al., 2018, Scheinost et al., 2019, Marek et al., 2022). The lower prediction accuracies associated with larger sample sizes are often due to the more population-representative samples. They are more accurate representations of the achievable prediction performance in practical applications, and are hence more useful for future studies to refer to. The use of multiple cohorts comprising different age range, ethnicity (Li et al., 2022), and socioeconomic backgrounds will further improve the generalizability of prediction models derived. Some studies with excitingly high prediction accuracies tend to be used as benchmarks that new prediction models need to be compared against. However, we argue that models trained on multiple large cohorts covering different sample characteristics and accounted for confounding factors properly should be preferable as benchmarks, despite the potentially lower accuracies.

As the ultimate goals of CBPP models are practical application and scientific discovery, the replicability of discoveries and model generalizability may be as important as the prediction accuracies. From the model training perspective, these goals may be best approached by the use of multiple large datasets, potentially covering similar demographics as the whole target population for the practical application. From the evaluation perspective, external validation is crucial in proving the usefulness of the model.

5.7 The Potential and Pitfalls of Interpretable Models

We have demonstrated that an interpretable CBPP framework can be helpful for both feature selection and discovering neurobiological insights. By providing data-driven feature selection, interpretable models can be used in conjunction with more complex but hard-to-interpret models to garner both interpretability and predictive power. The insights provided by the model interpretation can be validated against and/or help to validate biomarkers, activation maps, brain-wide associations, etc. Furthermore, the interpretable region-wise framework allowed us to assess replicability and generalizability in terms of the brain-behavior association derived from the models, instead of only in terms of prediction accuracies. We believe that interpretable models are useful and important for future developments in psychometric prediction.

Nevertheless, the region-wise approach is limited by the same general limitations faced by the whole field. Without sufficient predictive power and generalizability, models may not be considered worth interpreting in the first place. The prediction accuracy of

the model affects not just the model interpretation based on the region-wise approach, but also other types of model interpretation methods. For instance, regression weights can be transformed to interpretable activation values using the Haufe transformation (Haufe et al., 2014), allowing model interpretation based on each connectivity between each pair of brain regions. However, regression weights assigned by the prediction model can only be informative of the underlying brain-behavior relationships if the model is predictive enough. In study 2, we also demonstrated that the prediction patterns formed by the Haufe transformed weights are not replicable across cohorts, and tend to be affected by cohort-specific idiosyncrasies.

Alternatively, feature selection can be done by picking features that are significantly correlated with the target psychometric variable in the training set. Across cross-validation splits, features that are consistently selected may be considered related to the target psychometric measure (Jiang et al., 2018, Jiang et al., 2020). It would seem that these features' relationships with the target psychometric measure are not reliant on the prediction model, which is built after feature selection. Nevertheless, the statistical significance tests on the correlation between features and psychometric variables tend to be lenient. In the case where no real relationships exist between the neuroimaging features and the psychometric measure, it may still be possible to select a set of features that are lowly associated to the psychometric measure due to noise. In the end, the validity of the brain-behavior associations represented by these selected features would still depend on the predictive power of the prediction model. As it has also been shown that this method of feature selection can be unstable across cross-validation splits (O'Connor et al., 2021), it is unclear how generalizable these selected features would be.

5.8 Conclusion

In this work, we have developed and validated a region-wise CBPP framework for improving the interpretability of psychometric prediction models. We demonstrated two applications of the framework for constructing individual brain region's psychometric prediction profiles and for representing the prediction pattern representing the derived brain-behavior relationships for a psychometric variable. We then demonstrated the usefulness of the region-wise framework in examining cross-cohort replicability and generalizability of CBPP models. Finally, we presented the trends and issues in current

psychometric prediction studies, with a recommendation on the use of larger sample size and external validation. Overall, we hope to highlight the importance of model interpretability and generalizability, as well as the potentials of interpretable models. In order to reach the goal of practical applications, CBPP models still need to be improved in terms of predictive power, transparency, and generalizability.

Bibliography

Andersson, J.L., Jenkinson, M., Smith, S., et al. 2007. Non-linear registration, aka Spatial normalisation FMRIB technical report TR07JA2. *FMRIB Analysis Group of the University of Oxford*. 2(1), e21.

Anticevic, A., Dierker, D.L., Gillespie, S.K., Repovs, G., Csernansky, J.G., Van Essen, D.C., and Barch, D.M. 2008. Comparing surface-based and volume-based analyses of functional neuroimaging data in patients with schizophrenia. *Neuroimage*. 41(3), 835–848. DOI: 10.1016/j.neuroimage.2008.02.052.

Arbabshirani, M.R., Kiehl, K.A., Pearlson, G.D., and Calhoun, V.D. 2013. Classification of schizophrenia patients based on resting-state functional network connectivity. *Frontiers in Neuroscience*. 7, 133. DOI: 10.3389/fnins.2013.00133.

Avants, B.B., Tustison, N., Song, G., et al. 2009. Advanced normalization tools (ANTS). *Insight j*. 2(365), 1–35.

Beaty, R.E., Kenett, Y.N., Christensen, A.P., Rosenberg, M.D., Benedek, M., Chen, Q., Fink, A., Qiu, J., Kwapil, T.R., Kane, M.J., et al. 2018. Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences of the United States of America*. 115(5), 1087–1092. DOI: 10.1073/pnas.1713532115.

Bijsterbosch, J. and Beckmann, C. 2017. An introduction to resting state fMRI functional connectivity. Oxford University Press.

Biswal, B., Yetkin, F.Z., Haughton, V.M., and Hyde, J.S. 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*. 34(4), 537–541. DOI: 10.1002/mrm.1910340409.

Bressler, S.L. 1995. Large-scale cortical networks and cognition. *Brain Research. Brain Research Reviews*. 20(3), 288–304. DOI: 10.1016/0165-0173(94)00016-I.

Bzdok, D. and Ioannidis, J.P.A. 2019. Exploration, inference, and prediction in neuroscience and biomedicine. *Trends in Neurosciences*. 42(4), 251–262. DOI: 10.1016/j.tins.2019.02.001.

Bzdok, D., Varoquaux, G., and Steyerberg, E.W. 2021. Prediction, not association, paves the road to precision medicine. *JAMA psychiatry*. 78(2), 127–128. DOI: 10.1001/jamapsychiatry.2020.2549.

Chen, J., Müller, V.I., Dukart, J., Hoffstaedter, F., Baker, J.T., Holmes, A.J., Vatansever, D., Nickl-Jockschat, T., Liu, X., Derntl, B., et al. 2021. Intrinsic Connectivity Patterns of Task-Defined Brain Networks Allow Individual Prediction of Cognitive Symptom Dimension of Schizophrenia and Are Linked to Molecular Architecture. *Biological Psychiatry*. 89(3), 308–319. DOI: 10.1016/j.biopsych.2020.09.024.

Clos, M., Amunts, K., Laird, A.R., Fox, P.T., and Eickhoff, S.B. 2013. Tackling the multifunctional nature of Broca's region meta-analytically: co-activation-based parcellation of area 44. *Neuroimage*. 83, 174–188. DOI: 10.1016/j.neuroimage.2013.06.041.

Dadi, K., Rahim, M., Abraham, A., Chyzhyk, D., Milham, M., Thirion, B., Varoquaux, G., and Initiative, A.D.N. 2019. Benchmarking functional connectome-based predictive models for resting-state fMRI. *Neuroimage*. 192, 115–134. DOI: 10.1016/j.neuroimage.2019.02.062.

Dubois, J., Galdi, P., Han, Y., Paul, L.K., and Adolphs, R. 2018a. Resting-state functional brain connectivity best predicts the personality dimension of openness to experience. *Personality Neuroscience*. 1. DOI: 10.1017/pen.2018.8.

Dubois, J., Galdi, P., Paul, L.K., and Adolphs, R. 2018b. A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 373(1756). DOI: 10.1098/rstb.2017.0284.

Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., et al. 2019. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*. 16(1), 111–116. DOI: 10.1038/s41592-018-0235-4.

Ferguson, M.A., Anderson, J.S., and Spreng, R.N. 2017. Fluid and flexible minds: Intelligence reflects synchrony in the brain's intrinsic network architecture. *Network neuroscience (Cambridge, Mass.)* 1(2), 192–207. DOI: 10.1162/{NETN_a_00010}.

Finn, E.S. 2021. Is it time to put rest to rest? *Trends in Cognitive Sciences*. 25(12), 1021–1032. DOI: 10.1016/j.tics.2021.09.005.

Finn, E.S. and Bandettini, P.A. 2021. Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *Neuroimage*. 235, 117963. DOI: 10.1016/j.neuroimage.2021.117963.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., and Constable, R.T. 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience*. 18(11), 1664–1671. DOI: 10.1038/nn.4135.

Fox, M.D. and Raichle, M.E. 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews. Neuroscience*. 8(9), 700–711. DOI: 10.1038/nrn2201.

Fox, P.T. and Friston, K.J. 2012. Distributed processing; distributed functions? *Neuroimage*. 61(2), 407–426. DOI: 10.1016/j.neuroimage.2011.12.051.

Friston, K.J. 2011. Functional and effective connectivity: a review. *Brain connectivity*. 1(1), 13–36. DOI: 10.1089/brain.2011.0008.

Friston, K.J. 1994. Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*. 2(1-2), 56–78. DOI: 10.1002/hbm.460020107.

Gao, S., Greene, A.S., Constable, R.T., and Scheinost, D. 2019. Combining multiple connectomes improves predictive modeling of phenotypic measures. *Neuroimage*. 201, 116038. DOI: 10.1016/j.neuroimage.2019.116038.

Genon, S., Eickhoff, S., and Kharabian, S. 2022. Linking interindividual variability in brain structure to behaviour. *Nature Reviews. Neuroscience*. 23(5), 307–318. DOI: 10.1038/s41583-022-00584-7.

Genon, S., Reid, A., Langner, R., Amunts, K., and Eickhoff, S.B. 2018a. How to characterize the function of a brain region. *Trends in Cognitive Sciences*. 22(4), 350–364. ISSN: 13646613. DOI: 10.1016/j.tics.2018.01.010.

Genon, S., Reid, A., Li, H., Fan, L., Müller, V.I., Cieslik, E.C., Hoffstaedter, F., Langner, R., Grefkes, C., Laird, A.R., et al. 2018b. The heterogeneity of the left dorsal premotor cortex evidenced by multimodal connectivity-based parcellation and functional characterization. *Neuroimage*. 170, 400–411. DOI: 10.1016/j.neuroimage.2017.02.034.

Ghosh, S.S., Kakunoori, S., Augustinack, J., Nieto-Castanon, A., Kovelman, I., Gaab, N., Christodoulou, J.A., Triantafyllou, C., Gabrieli, J.D.E., and Fischl, B. 2010. Evaluating the validity of volume-based and surface-based brain image

registration for developmental cognitive neuroscience studies in children 4 to 11 years of age. *Neuroimage*. 53(1), 85–93. DOI: 10.1016/j.neuroimage.2010.05.075.

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., et al. 2013. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage*. 80, 105–124. DOI: 10.1016/j.neuroimage.2013.04.127.

Greene, A.S., Gao, S., Scheinost, D., and Constable, R.T. 2018. Task-induced brain state manipulation improves prediction of individual traits. *Nature Communications*. 9(1), 2807. DOI: 10.1038/s41467-018-04920-3.

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., Zsoldos, E., Ebmeier, K.P., Filippini, N., Mackay, C.E., et al. 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage*. 95, 232–247. DOI: 10.1016/j.neuroimage.2014.03.034.

Hacker, C.D., Laumann, T.O., Szrama, N.P., Baldassarre, A., Snyder, A.Z., Leuthardt, E.C., and Corbetta, M. 2013. Resting state network estimation in individual subjects. *Neuroimage*. 82, 616–633. DOI: 10.1016/j.neuroimage.2013.05.108.

Hastie, T., Tibshirani, R., and Friedman, J. 2009. The Elements of Statistical Learning. 2nd. Springer Series in Statistics. New York, NY: Springer New York, 106–119. DOI: 10.1007/978-0-387-84858-7.

Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., and Bie, F. 2014. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage*. 87, 96–110. DOI: 10.1016/j.neuroimage.2013.10.067.

He, T., Kong, R., Holmes, A.J., Nguyen, M., Sabuncu, M.R., Eickhoff, S.B., Bzdok, D., Feng, J., and Yeo, B.T.T. 2020. Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *Neuroimage*. 206, 116276. DOI: 10.1016/j.neuroimage.2019.116276.

Jain, A. and Waller, W. 1978. On the optimal number of features in the classification of multivariate Gaussian data. *Pattern recognition*. 10(5-6), 365–374. DOI: 10.1016/0031-3203(78)90008-0.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*. 17(2), 825–841. DOI: 10.1006/nimg.2002.1132.

Jiang, R., Calhoun, V.D., Fan, L., Zuo, N., Jung, R., Qi, S., Lin, D., Li, J., Zhuo, C., Song, M., et al. 2020. Gender Differences in Connectome-based Predictions of Individualized

Intelligence Quotient and Sub-domain Scores. *Cerebral Cortex*. 30(3), 888–900. DOI: 10.1093/cercor/bhz134.

Jiang, R., Calhoun, V.D., Zuo, N., Lin, D., Li, J., Fan, L., Qi, S., Sun, H., Fu, Z., Song, M., et al. 2018. Connectome-based individualized prediction of temperament trait scores. *Neuroimage*. 183, 366–374. DOI: 10.1016/j.neuroimage.2018.08.038.

Karcher, N.R. and Barch, D.M. 2021. The ABCD study: understanding the development of risk for mental and physical health outcomes. *Neuropsychopharmacology*. 46(1), 131–142. DOI: 10.1038/s41386-020-0736-6.

Karolis, V.R., Corbetta, M., and Thiebaut de Schotten, M. 2019. The architecture of functional lateralisation and its relationship to callosal connectivity in the human brain. *Nature Communications*. 10(1), 1417. DOI: 10.1038/s41467-019-09344-1.

Kong, R., Li, J., Orban, C., Sabuncu, M.R., Liu, H., Schaefer, A., Sun, N., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., et al. 2019. Spatial Topography of Individual-Specific Cortical Networks Predicts Human Cognition, Personality, and Emotion. *Cerebral Cortex*. 29(6), 2533–2551. DOI: 10.1093/cercor/bhy123.

Kong, R., Yang, Q., Gordon, E., Xue, A., Yan, X., Orban, C., Zuo, X.-N., Spreng, N., Ge, T., Holmes, A., et al. 2021. Individual-Specific Areal-Level Parcellations Improve Functional Connectivity Prediction of Behavior. *Cerebral Cortex*. 31(10), 4477–4500. DOI: 10.1093/cercor/bhab101.

Lemieux, L., Salek-Haddadi, A., Lund, T.E., Laufs, H., and Carmichael, D. 2007. Modelling large motion events in fMRI studies of patients with epilepsy. *Magnetic Resonance Imaging*. 25(6), 894–901. DOI: 10.1016/j.mri.2007.03.009.

Li, J., Bzdok, D., Chen, J., Tam, A., Ooi, L.Q.R., Holmes, A.J., Ge, T., Patil, K.R., Jabbi, M., Eickhoff, S.B., et al. 2022. Cross-ethnicity/race generalization failure of behavioral prediction from resting-state functional connectivity. *Science Advances*. 8(11), eabj1812. DOI: 10.1126/sciadv.abj1812.

Li, J., Kong, R., Liégeois, R., Orban, C., Tan, Y., Sun, N., Holmes, A.J., Sabuncu, M.R., Ge, T., and Yeo, B.T.T. 2019. Global signal regression strengthens association between resting-state functional connectivity and behavior. *Neuroimage*. 196, 126–141. DOI: 10.1016/j.neuroimage.2019.04.016.

Lim, S., Radicchi, F., Heuvel, M.P. van den, and Sporns, O. 2019. Discordant attributes of structural and functional brain connectivity in a two-layer multiplex network. *Scientific Reports*. 9(1), 2885. DOI: 10.1038/s41598-019-39243-w.

Maglanoc, L.A., Kaufmann, T., Meer, D. van der, Marquand, A.F., Wolfers, T., Jonassen, R., Hilland, E., Andreassen, O.A., Landrø, N.I., and Westlye, L.T. 2020. Brain connectome mapping of complex human traits and their polygenic architecture using machine learning. *Biological Psychiatry*. 87(8), 717–726. DOI: 10.1016/j.biopsych.2019.10.011.

Marek, S., Tervo-Clemmens, B., Calabro, F.J., Montez, D.F., Kay, B.P., Hatoum, A.S., Donohue, M.R., Foran, W., Miller, R.L., Hendrickson, T.J., et al. 2022. Reproducible brain-wide association studies require thousands of individuals. *Nature*. 603(7902), 654–660. DOI: 10.1038/s41586-022-04492-9. (Visited on 07/13/2022).

Miller, K.M., Price, C.C., Okun, M.S., Montijo, H., and Bowers, D. 2009. Is the n-back task a valid neuropsychological measure for assessing working memory? *Archives of Clinical Neuropsychology*. 24(7), 711–717. DOI: 10.1093/arclin/acp063.

Mueller, S., Wang, D., Fox, M.D., Yeo, B.T.T., Sepulcre, J., Sabuncu, M.R., Shafee, R., Lu, J., and Liu, H. 2013. Individual variability in functional connectivity architecture of the human brain. *Neuron*. 77(3), 586–595. DOI: 10.1016/j.neuron.2012.12.028.

Nakamura, K., Kawashima, R., Sato, N., Nakamura, A., Sugiura, M., Kato, T., Hatano, K., Ito, K., Fukuda, H., Schormann, T., et al. 2000. Functional delineation of the human occipito-temporal areas related to face and scene processing. A PET study. *Brain: A Journal of Neurology*. 123 (Pt 9), 1903–1912. DOI: 10.1093/brain/123.9.1903.

Nelson, C.A. 2001. The development and neural bases of face recognition. *Infant and child development*. 10(1-2), 3–18. DOI: 10.1002/icd.239.

Nickerson, L.D., Smith, S.M., Öngür, D., and Beckmann, C.F. 2017. Using dual regression to investigate network shape and amplitude in functional connectivity analyses. *Frontiers in Neuroscience*. 11, 115. DOI: 10.3389/fnins.2017.00115.

Nooner, K.B., Colcombe, S.J., Tobe, R.H., Mennes, M., Benedict, M.M., Moreno, A.L., Panek, L.J., Brown, S., Zavitz, S.T., Li, Q., et al. 2012. The NKI-Rockland Sample: A Model for Accelerating the Pace of Discovery Science in Psychiatry. *Frontiers in Neuroscience*. 6, 152. DOI: 10.3389/fnins.2012.00152.

Nostro, A.D., Müller, V.I., Varikuti, D.P., Pläschke, R.N., Hoffstaedter, F., Langner, R., Patil, K.R., and Eickhoff, S.B. 2018. Predicting personality from network-based resting-state functional connectivity. *Brain Structure & Function*. 223(6), 2699–2719. DOI: 10.1007/s00429-018-1651-z.

O'Connor, D., Lake, E.M.R., Scheinost, D., and Constable, R.T. 2021. Resample aggregating improves the generalizability of Connectome Predictive Modelling. *Neuroimage*, 118044. DOI: 10.1016/j.neuroimage.2021.118044.

Ombao, H. 2016. Handbook of neuroimaging data analysis. Chapman and Hall/CRC. DOI: 10.1201/9781315373652.

Pan, G., Xiao, L., Bai, Y., Wilson, T.W., Stephen, J.M., Calhoun, V.D., and Wang, Y.-P. 2021. Multiview Diffusion Map Improves Prediction of Fluid Intelligence With Two Paradigms of fMRI Analysis. *IEEE Transactions on Bio-Medical Engineering*. 68(8), 2529–2539. DOI: 10.1109/{TBME}.2020.3048594.

Pervaiz, U., Vidaurre, D., Woolrich, M.W., and Smith, S.M. 2020. Optimising network modelling methods for fMRI. *Neuroimage*. 211, 116604. DOI: 10.1016/j.neuroimage.2020.116604.

Pläschke, R.N., Patil, K.R., Cieslik, E.C., Nostro, A.D., Varikuti, D.P., Plachti, A., Lösche, P., Hoffstaedter, F., Kalenscher, T., Langner, R., et al. 2020. Age differences in predicting working memory performance from network-based functional connectivity. *Cortex*. 132, 441–459. DOI: 10.1016/j.cortex.2020.08.012.

Poldrack, R.A., Huckins, G., and Varoquaux, G. 2020. Establishment of best practices for evidence for prediction: A review. *JAMA psychiatry*. 77(5), 534–540. DOI: 10.1001/jamapsychiatry.2019.3671.

Poldrack, R.A., Nichols, T., and Mumford, J. 2011. Handbook of functional MRI data analysis. Cambridge University Press. DOI: 10.1017/{CBO9780511895029}.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., and Petersen, S.E. 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*. 59(3), 2142–2154. DOI: 10.1016/j.neuroimage.2011.10.018.

Power, J.D., Plitt, M., Laumann, T.O., and Martin, A. 2017. Sources and implications of whole-brain fMRI signals in humans. *Neuroimage*. 146, 609–625. DOI: 10.1016/j.neuroimage.2016.09.038.

Pruim, R.H.R., Mennes, M., Buitelaar, J.K., and Beckmann, C.F. 2015a. Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. *Neuroimage*. 112, 278–287. DOI: 10.1016/j.neuroimage.2015.02.063.

Pruim, R.H.R., Mennes, M., Rooij, D. van, Llera, A., Buitelaar, J.K., and Beckmann, C.F. 2015b. ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage*. 112, 267–277. DOI: 10.1016/j.neuroimage.2015.02.064.

Rasero, J., Sentis, A.I., Yeh, F.-C., and Verstynen, T. 2021. Integrating across neuroimaging modalities boosts prediction accuracy of cognitive ability. *PLoS Computational Biology*. 17(3), e1008347. DOI: 10.1371/journal.pcbi.1008347.

Rogalsky, C. and Hickok, G. 2011. The role of Broca's area in sentence comprehension. *Journal of Cognitive Neuroscience*. 23(7), 1664–1680. DOI: 10.1162/jocn.2010.21530.

Rosenberg, M.D., Finn, E.S., Scheinost, D., Papademetris, X., Shen, X., Constable, R.T., and Chun, M.M. 2016. A neuromarker of sustained attention from whole-brain functional connectivity. *Nature Neuroscience*. 19(1), 165–171. DOI: 10.1038/nn.4179.

Salmon, E., Van der Linden, M., Collette, F., Delfiore, G., Maquet, P., Degueldre, C., Luxen, A., and Franck, G. 1996. Regional brain activity during working memory tasks. *Brain: A Journal of Neurology*. 119 (Pt 5), 1617–1625. DOI: 10.1093/brain/119.5.1617.

Sams, M., Hietanen, J.K., Hari, R., Ilmoniemi, R.J., and Lounasmaa, O.V. 1997. Face-specific responses from the human inferior occipito-temporal cortex. *Neuroscience*. 77(1), 49–55. DOI: 10.1016/s0306-4522(96)00419-8.

Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., et al. 2013. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage*. 64, 240–256. DOI: 10.1016/j.neuroimage.2012.08.052.

Scheinost, D., Noble, S., Horien, C., Greene, A.S., Lake, E.M., Salehi, M., Gao, S., Shen, X., O'Connor, D., Barron, D.S., et al. 2019. Ten simple rules for predictive modeling of individual differences in neuroimaging. *Neuroimage*. 193, 35–45. DOI: 10.1016/j.neuroimage.2019.02.057.

Scheinost, D., Stoica, T., Wasylink, S., Gruner, P., Saksa, J., Pittenger, C., and Hampson, M. 2014. Resting state functional connectivity predicts neurofeedback response. *Frontiers in Behavioral Neuroscience*. 8, 338. DOI: 10.3389/fnbeh.2014.00338.

Sen, B. and Parhi, K.K. 2021. Predicting Biological Gender and Intelligence From fMRI via Dynamic Functional Connectivity. *IEEE Transactions on Bio-Medical Engineering*. 68(3), 815–825. DOI: 10.1109/{TBME}.2020.3011363.

Smith, E.E. and Jonides, J. 1999. Storage and executive processes in the frontal lobes. *Science*. 283(5408), 1657–1661. DOI: 10.1126/science.283.5408.1657.

Soares, J.M., Magalhães, R., Moreira, P.S., Sousa, A., Ganz, E., Sampaio, A., Alves, V., Marques, P., and Sousa, N. 2016. A hitchhiker's guide to functional magnetic resonance imaging. *Frontiers in Neuroscience*. 10, 515. DOI: 10.3389/fnins.2016.00515.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. 2015. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Medicine*. 12(3), e1001779. DOI: 10.1371/journal.pmed.1001779.

Sui, J., Jiang, R., Bustillo, J., and Calhoun, V. 2020. Neuroimaging-based Individualized Prediction of Cognition and Behavior for Mental Disorders and Health: Methods and Promises. *Biological Psychiatry*. 88(11), 818–828. DOI: 10.1016/j.biopsych.2020.02.016.

Tononi, G., Edelman, G.M., and Sporns, O. 1998. Complexity and coherency: integrating information in the brain. *Trends in Cognitive Sciences*. 2(12), 474–484. DOI: 10.1016/s1364-6613(98)01259-5.

Tononi, G., Sporns, O., and Edelman, G.M. 1994. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences of the United States of America*. 91(11), 5033–5037. DOI: 10.1073/pnas.91.11.5033.

Tucholka, A., Fritsch, V., Poline, J.-B., and Thirion, B. 2012. An empirical comparison of surface-based and volume-based group studies in neuroimaging. *Neuroimage*. 63(3), 1443–1453. DOI: 10.1016/j.neuroimage.2012.06.019.

Turner, B.O., Paul, E.J., Miller, M.B., and Barbey, A.K. 2018. Small sample sizes reduce the replicability of task-based fMRI studies. *Communications Biology*. 1, 62. DOI: 10.1038/s42003-018-0073-z.

Van Essen, D.C., Glasser, M.F., Dierker, D.L., Harwell, J., and Coalson, T. 2012. Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cerebral Cortex*. 22(10), 2241–2262. DOI: 10.1093/cercor/bhr291.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., and Consortium, W.-M.H. 2013. The WU-Minn Human Connectome Project: an overview. *Neuroimage*. 80, 62–79. DOI: 10.1016/j.neuroimage.2013.05.041.

Wu, J., Eickhoff, S.B., Hoffstaedter, F., Patil, K.R., Schwender, H., Yeo, B.T.T., and Genon, S. 2021. A Connectivity-Based Psychometric Prediction Framework for Brain-Behavior Relationship Studies. *Cerebral Cortex*. 31(8), 3732–3751. DOI: 10.1093/cercor/bhab044.

Wu, J., Li, J., Eickhoff, S.B., Hoffstaedter, F., Hanke, M., Yeo, B.T., and Genon, S. 2022. Cross-cohort replicability and generalizability of connectivity-based psychometric prediction patterns. *Neuroimage*, 119569. DOI: 10.1016/j.neuroimage.2022.119569.

Yeung, A.W.K., More, S., Wu, J., and Eickhoff, S.B. 2022. Reporting details of neuroimaging studies on individual traits prediction: a literature survey. *Neuroimage*, 119275. DOI: 10.1016/j.neuroimage.2022.119275.

Zou, H. and Hastie, T. 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 67(2), 301–320. DOI: 10.1111/j.1467-9868.2005.00503.x.

Zurowski, B., Gostomzyk, J., Grön, G., Weller, R., Schirrmeister, H., Neumeier, B., Spitzer, M., Reske, S.N., and Walter, H. 2002. Dissociating a common working memory network from different neural substrates of phonological and spatial stimulus processing. *Neuroimage*. 15(1), 45–57. DOI: 10.1006/nimg.2001.0968.

Acknowledgements

I would like to thank my supervisors, Professor Sarah Genon, Professor Simon B. Eichoff, and Professor Holger Schwender for their instrumental support and guidance. They have helped me to grow into an independent researcher ready for future career in science. Further thanks go to my colleague and collaborators, Jingwei Li, Felix Hoffstaedter, Eliana Nicholaisen, Shammi More, Professor Michael Hanke, and Laura Waite for providing valuable helps in making my work possible. I would also like to express my gratitude to Professor B.T. Thomas Yeo, who has inducted me into the neuroimaging community and introduced me the essential skills of scientific research. A very special thank goes to Niels Reuter who has filled my life with joy, appreciation, and care. Finally, I am and will always be grateful to my parents for their continued support and understanding.