NeuroImage 276 (2023) 120205

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Decoding individual differences in self-prioritization from the resting-state functional connectome

Yongfa Zhang [a], Fei Wang [a,b,c,*], Jie Sui [d]

[a] Department of Psychology, School of Social Sciences, Tsinghua University, Beijing 100084, China
[b] Laboratory of Brain and Intelligence, Tsinghua University, Beijing 100084, China
[c] The Centre for Positive Psychology Research, Tsinghua University, Beijing 100084, China
[d] School of Psychology, University of Aberdeen, Aberdeen AB24 3FX, Scotland, Great Britain

ARTICLE INFO

Keywords:
Self-prioritization effect
Resting state
fMRI
Functional connectivity
Machine learning

ABSTRACT

Although the self has traditionally been viewed as a higher-order mental function by most theoretical frameworks, recent research advocates a *fundamental self hypothesis*, viewing the self as a baseline function of the brain embedded within its spontaneous activities, which dynamically regulates cognitive processing and subsequently guides behavior. Understanding this fundamental self hypothesis can reveal where self-biased behaviors emerge and to what extent brain signals at rest can predict such biased behaviors. To test this hypothesis, we investigated the association between spontaneous neural connectivity and robust self-bias in a perceptual matching task using resting-state functional magnetic resonance imaging (fMRI) in 348 young participants. By decoding whole-brain connectivity patterns, the support vector regression model produced the best predictions of the magnitude of self-bias in behavior, which was evaluated via a nested cross-validation procedure. The out-of-sample generalizability was further authenticated using an external dataset of older adults. The functional connectivity results demonstrated that self-biased behavior was associated with distinct connections between the default mode, cognitive control, and salience networks. Consensus network and computational lesion analyses further revealed contributing regions distributed across six networks, extending to additional nodes, such as the thalamus, whose role in self-related processing remained unclear. These results provide evidence that self-biased behavior derives from spontaneous neural connectivity, supporting the fundamental self hypothesis. Thus, we propose an integrated neural network model of this fundamental self that synthesizes previous theoretical models and portrays the brain mechanisms by which the self emerges at rest internally and regulates responses to the external environment.

1. Introduction

One of the most intriguing questions in science is how the brain generates a sense of the self. Previous philosophical, psychological, and neuroscientific discussions on this topic have yielded two competing views (Northoff, 2016). The first, hereafter referred to as the *higher-order self hypothesis*, dates back to philosophers such as Descartes and Kant and regards the self as one of the brain's highest functions and serves as a meta-representation of other lower-level functions such as perception, emotion, and action (Damasio, 1999; Northoff, 2012). However, recent research has challenged this notion and advocated for an alternative *fundamental self hypothesis*, which depicts the self as a fundamental brain function that precedes and regulates other functions (Humphreys and Sui, 2016; Northoff, 2016; Northoff et al., 2022; Qin et al., 2020; Sui and Gu, 2017; Sui and Humphreys, 2015). According to this viewpoint, the self is an intrinsic feature of the brain that is embedded within and continuously sustained in spontaneous activities of the brain (Northoff, 2016). When an internal or external stimulus appears, the self network interacts with task-related networks to exert its influence on cognitive and affective processing (Qin et al., 2020; Sui and Gu, 2017).

Northoff (2016) initially proposed the notion of the fundamental self, known as a basis model of self-specificity (BMSS), which argues for a relationship between spontaneous brain activity and basic human functions such as perception and emotions. Here, the term "fundamental" does not imply that the self is the most basic function in the brain (nor could it be) but rather its independence and priority from other basic functions. In this context, the self refers to a moment-by-moment spontaneous function that intrinsically integrates bodily and mental aspects (Qin et al., 2020). Sui and Gu (2017) summarized recent evidence regarding brain networks supporting the fundamental self found in various experimental paradigms in healthy participants and those with psychi-

* Corresponding author.
E-mail address: wf3126@mail.tsinghua.edu.cn (F. Wang).

https://doi.org/10.1016/j.neuroimage.2023.120205.
Received 13 January 2023; Received in revised form 23 May 2023; Accepted 27 May 2023
Available online 28 May 2023.

atric disorders, such as the perceptual matching task and self in embodied and interoceptive tasks. They proposed a *self-as-object* model that explains how three neural networks—the default mode network, cognitive control network, and the salience network—interact to control behavior in the presence of self-related stimuli. Similarly, Qin et al. (2020) conducted a meta-analysis to examine a three-layer neural model of self, including Interoceptive-processing, Exteroceptive-processing, and Mental-self processing, which explains how external stimuli become self-related and thereby integrated within the self.

Although these previous studies and tasks have been developed to address specific aspects of the self, neural networks supporting the fundamental self have been identified from an integrative perspective (Northoff, 2016; Qin et al., 2020; Sui and Gu, 2017). However, there is little direct evidence on whether the fundamental self at the whole-brain level can predict self-biased behavior. This study was set out to address this issue by examining whether the resting-state connectome can predict self-biased behavior, focusing on a common psychological finding—the *self-prioritization effect* (SPE) in perceptual matching—which suggests that self-related information is prioritized during information processing (Sui et al., 2012).

It has been well documented that our brain prioritizes processing stimuli related to ourselves (Kotlewska and Nowicka, 2015, 2016; Macrae et al., 2018; Sui and Humphreys, 2017b). The SPE manifests itself in a variety of cognitive processes, including memory, attention, and perception (for an overview, see Cunningham and Turk, 2017). For example, individuals can automatically notice the mention of their name in background noise (Moray, 1959) and recognize their own faces more quickly than other people's faces (Humphreys and Sui, 2016; Tacikowski and Nowicka, 2010; Wójcik et al., 2019). Even temporarily tagging neutral stimuli, such as geometric shapes, colors, and gratings, to the self in a few trials can generate a robust SPE to these stimuli associated with the self (Sui et al., 2012). Given the widespread manifestation of the SPE at various levels of cognitive processing, there might be a link between intrinsic brain networks at the whole brain level and self-processing.

According to the BMSS (Northoff, 2016), if self-specificity is a fundamental component of spontaneous brain activity that can easily integrate external stimuli with the self, there should be a distinctive pattern of the brain's resting-state networks (RSN) to self-biased behavior. Brain regions supporting this idea have been identified as cortical midline structures (Northoff and Bermpohl, 2004). The default mode network (DMN), which is the most well-established RSN and is primarily active during rest, overlaps with the network activated when one thinks about oneself, such as the medial prefrontal cortex (MPFC) (Davey et al., 2016; Northoff and Bermpohl, 2004; Qin and Northoff, 2011). In addition, it has been reported that the processing of self-related stimuli influences subsequent resting-state activity and vice versa (Qin et al., 2016; Schneider et al., 2008; Wang et al., 2013), and that resting-state activities can be used to predict self-related characteristics such as self-consciousness (Huang et al., 2016; Wolff et al., 2019) or self-reflectiveness (Larabi et al., 2020). This evidence suggests that the RSN, especially the DMN, might be able to predict self-biased behavior such as the SPE.

In addition, self functions are closely associated with basic psychological processes of emotion and reward, as shown by the involvement of the limbic and reward systems in self-related tasks (de Greck et al., 2008; Kircher et al., 2000; Northoff, 2016). This indicates that not only is it important to identify cortical brain connections, but recent evidence suggests that the contribution of the subcortical structures in the fundamental self model should be evaluated. Meta-analyses have demonstrated the role of several subcortical structures in self-processing, such as the thalamus (Murray et al., 2015; Qin et al., 2020). This is consistent with developmental evidence indicating that newborn infants have a rudimentary capacity to distinguish between self and non-self and that their explicit self-awareness typically emerges around the second year of life (for reviews, see Gallagher and Meltzoff, 1996; Rochat, 2003). As a result, we expected brain functional connections between cortical and subcortical regions to predict self-biased behavior.

In summary, we explored the fundamental self hypothesis in a large sample of young and healthy participants (n = 348) by predicting the SPE in perceptual matching (Sui et al., 2012) from the resting-state connectome. As the SPE in perceptual shows the stability of self-processing across contexts and tasks (Scheller and Sui, 2022) and is associated with different aspects of cognitive processing (Sui and Humphreys, 2015), it may be an appropriate tool for examining how self-specificity regulates mental functions, especially low-level functions such as attention and perception, and thereby it may serve as a suitable test of the fundamental self hypothesis (Sui and Gu, 2017; Sui et al., 2012; Sui and Humphreys, 2015). In addition, in contrast to previous studies' univariate general linear model approach that treated individual differences as noise (Scheinost et al., 2019; Sui et al., 2013), we used a machine learning approach by building a linear support vector regression (SVR) model with whole-brain resting-state functional connectivity (rs-FC) as the input feature and the predicted magnitude of the SPE as the output. This approach leverages rich information on the spontaneous brain activity on rs-fMRI and the data-driven nature of machine learning to generate individualized predictions. It has been used successfully to identify neurobiological markers of individual difference variables such as demographics (Zhang et al., 2018), traits (Jiang et al., 2018), and performance in cognitive tasks (Rosenberg et al., 2016; Sripada et al., 2020). A rigorous leave-one-out cross-validation (LOOCV) procedure was applied to ensure out-of-sample generalizability (Magnusson et al., 2019). The contributing FCs were further identified through analyses of model coefficients and computational lesions.

2. Methods

2.1. Participants

A total of 380 healthy young participants were recruited to carry out a number of questionnaires and behavioral tasks while resting state and brain structure scans were conducted. After excluding those who failed to complete the perceptual matching task or lacked demographic data (n = 20), or had excessive head motion during scanning (n = 12, defined as > 0.2 mm mean framewise displacement), the final sample used for analysis consisted of 348 individuals (182 women; mean age, 22.5 years; age range, 18–34 years). All the participants provided informed consent. This study was conducted under an ethical protocol approved by the Ethics Committee of the Center for Biomedical Imaging Research at Tsinghua University.

2.2. Behavioral assessment

We used a well-established perceptual matching paradigm to quantify the magnitude of an individual's self-prioritization (Fig. 1a; Sui et al., 2012). This paradigm has not only overcome the familiarity advantage of self-relevant stimuli (e.g., own name vs. others' name), but has also shown resistance to confounding factors such as word length (Sui et al., 2012; Woźniak and Knoblich, 2019), and thus has been widely adopted by the academic community (Yankouskaya and Sui, 2022). There were two phases in this task. In the learning phase, participants were instructed to associate two geometric shapes (i.e., triangle and circle) with two personalized labels (i.e., "you" and "stranger"). For example, participants may be told that a circle represents themselves and a triangle represents a stranger. The matching pattern was counterbalanced across participants, who were asked to judge whether an upcoming series of random shape-label pairs corresponded to it. All instructions were presented on the first screen of the experimental program until participants made a response indicating that they were ready to begin the experiment.

After the associations were created, participants started the matching phase immediately. Each trial started with a central fixation cross

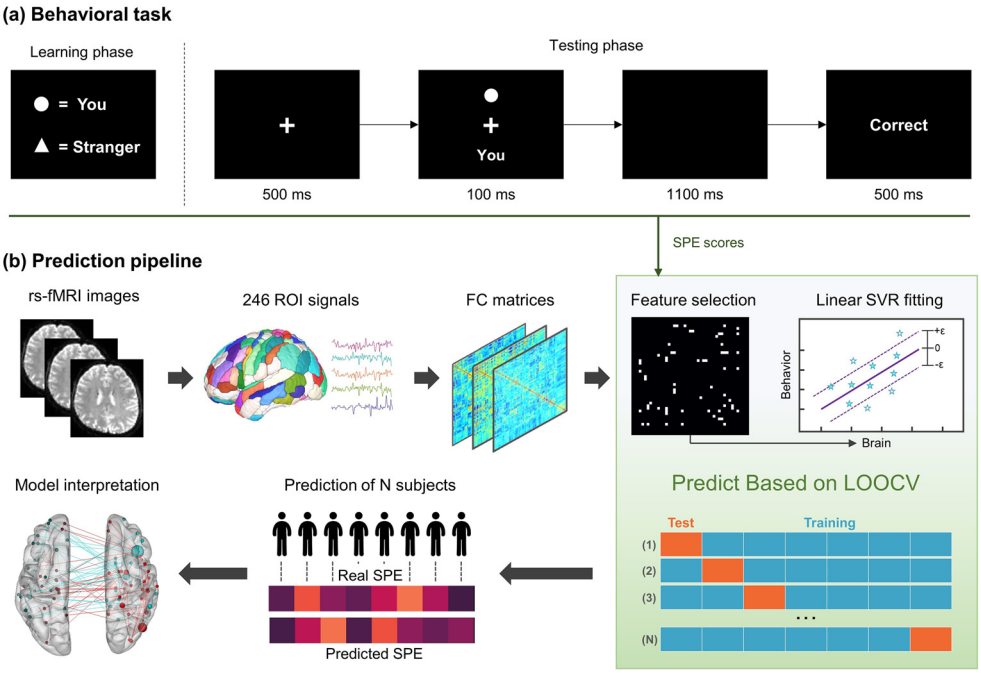

Fig. 1. Overview of our approach. (a) Example of stimuli and procedures of the perceptual matching task. (b) Schematic flow of prediction pipeline. SPE, self-prioritization effect; rs-fMRI, resting-state functional magnetic resonance imaging; ROI, region of interest; FC, functional connectivity; SVR, support vector regression; LOOCV, leave-one-out cross-validation.

(0.8° × 0.8° visual angle) for 500 ms. Then, a shape (3.8° × 3.8°) and a label (2.4° × 1.6°) were presented above and below the fixation cross for 100 ms. A subsequent blank frame occupied the screen until a response was made or 1100 ms elapsed, during which the participants were expected to judge whether the shape-label pairing was correct as quickly and accurately as possible. Feedback (correct or incorrect) was provided on the screen for 500 ms at the end of each trial. Each participant completed 96 trials following 12 practice trials, with a total of 24 trials for each condition (self-matching, self-mismatching, stranger-matching, and stranger-mismatching). All stimuli were rendered on a gray background on a 17-inch monitor (1024 × 768 at 60 Hz). The experiment was run on a PC using E-prime software (version 2.0).

Previous studies using this paradigm have generally performed group-wise analyses to compare differences between the self and other conditions (e.g., Humphreys and Sui, 2015; Sui et al., 2013). However, machine learning techniques entail specific individualized scores for prediction. We defined the SPE score as the difference in mean reaction time (RT) between two matching conditions (i.e., SPE score = $RT_{stranger\text{-}matching} - RT_{self\text{-}matching}$) for the following reasons: 1) self-prioritization has been most robustly characterized by differences between the reaction time of the self and others in matching conditions (Sui et al., 2013); 2) the accuracy performance in these two conditions showed a ceiling effect with an average accuracy of over 91%, causing the accuracy distribution not to be suitable for individual-level analysis. This adopted metric has also been employed in previous neuroscience research on the SPE (e.g., Jiang and Sui, 2022; Stolte et al., 2017).

2.3. Image acquisition

Images were acquired with a Philips Achieva 3.0T TX system at the Center of Bio-Medical Imaging Research, Tsinghua University. All participants completed a 508.3-second rs-fMRI scanning, during which they were instructed to open their eyes, to not think about anything systematically, and to not fall asleep. Functional images were acquired with a T2-weighted echo-planar imaging sequence using the following parameters: repetition time (TR) = 2300 ms, echo time (TE) = 30 ms, flip angle = 90°, 37 ascending slices, field of view (FOV) = 256 × 256, acquisition matrix = 96 × 96 × 37, voxel size = 2.5 × 2.5 × 3.45 mm. A high-resolution T1-weighted structural image was also obtained for each participant using the following parameters: TR = 8.2 ms, TE = 3.8 ms, 160 contiguous sagittal slices, flip angle = 8°, FOV = 256 × 256, acquisition matrix = 256 × 256 × 160, voxel size = 0.938 × 0.938 × 1 mm; sensitivity encoding (SENSE) factor: anterior-posterior (AP) = 2, right-left (RL) = 1.5. Participants completed the behavioral assessment outside the MRI scanner.

2.4. Image preprocessing

Imaging data were preprocessed using the DPABI software package (version 6.0; Yan et al., 2016), which is based on Statistical Parametric Mapping (SPM12, https://www.fil.ion.ucl.ac.uk/spm/) and has been widely used. The preprocessing procedure included the following steps: (1) the first 10 functional volumes were removed to ensure signal stability; (2) the remaining 211 volumes were slice-time corrected and realigned to the first image; (3) the T1-weighted structural images were co-registered to the mean functional image and then segmented into gray matter, white matter, and cerebrospinal fluid; (4) the nuisance signals were regressed out, including the Friston 24 head-motion parameters (Friston et al., 1996), five principal components of white matter and cerebrospinal fluid signals, global signals, and linear signal trends; (5) derived images were normalized to the Montreal Neurological Institute space using the Diffeomorphic Anatomical Registration Through Exponentiated Lie algebra (DARTEL) algorithm and resampled to 3-mm cubic voxels; (6) the normalized images were spatially smoothed with a 4-mm Gaussian kernel and temporally smoothed using the frequency bandwidth of 0.01–0.1 Hz; (7) to further eliminate effects of head movement, as suggested by Power et al. (2012), we applied a common threshold to scrub the volumes with framewise diffusion greater than 0.5 mm (see also Sripada et al., 2020; Y. Wang et al., 2021).

Notably, in line with many neuroimaging prediction works (e.g., Goldfarb et al., 2020; Yu et al., 2020), we employed a contentious step, global signal regression (GSR), in this study. The primary consideration behind this procedure is that GSR not only efficiently removes artifacts caused by head movement and physiological factors but also robustly improves behavioral prediction performance (Li et al., 2019). However, given that the neuroscience community has not reached a consensus on the issue of GSR (Murphy and Fox, 2017) and that recent studies have re-highlighted its importance (Scalabrini et al., 2020;

Zhang and Northoff, 2022), we also conducted an analysis without GSR.

2.5. Functional connectome construction

As shown in Fig. 1b, the whole-brain functional network nodes were defined using the Human Brainnetome Atlas (Fan et al., 2016), which consists of 210 cortical and 36 subcortical regions of interest (ROIs). This fine-grained atlas reliably integrates information from anatomical and functional connections and has been used in many neuroimaging studies to predict individual differences (Yeung et al., 2022). For each participant, time series within each node were computed by averaging the blood oxygenation level-dependent (BOLD) signal over all voxels in the ROIs. We then calculated the Pearson correlation for each pair of the 246 resulting time series to obtain a 246 × 246 symmetric FC matrix for each participant. Each element in the matrix characterizes the alleged connection (also called edge) strength between two nodes.

2.6. Individualized prediction

When combining neuroimaging and behavioral data in machine learning algorithms, an inescapable challenge is the potential curse of dimensionality caused by few samples but high dimensionality (Scheinost et al., 2019). Although numerous voxels in the brain were grouped into 246 ROIs, more than 30,000 unique features (i.e., edges) remained in the FC matrix. Two typical solutions are to retain only the most relevant features before feeding them into the model and to use regularization techniques in the model to mitigate the influence of non-predictive features (Ying, 2019). We developed a prediction framework integrating feature selection and a regularized linear SVR model to predict individuals' SPE scores based on whole-brain FC (see Fig. S1).

We implemented the linear SVR using the Scikit-learn library (Pedregosa et al., 2011). SVR is a powerful supervised machine-learning algorithm extensively used in neuroscience and psychology studies (Cui and Gong, 2018). When constructed based on a linear kernel function, SVR can be highly predictive and interpretive. Linear SVR uses a regularization hyperparameter C to determine the model complexity tradeoff and thus has a significant impact on model performance. To avoid data leakage during hyperparameter determination, we employed a nested cross-validation (CV) procedure, which consisted of an outer loop using a LOOCV and an inner loop using a 5-fold CV. As demonstrated by Vabalas et al. (2019), only a nested CV framework with completely separated training and testing data can guarantee unbiased performance estimates.

The inner 5-fold CV was applied to quickly find the best hyperparameters (i.e., C), a process also known as grid search. The dataset input to the inner loop was randomly divided into five parts, four of which were regarded as the subtraining set and the remaining part as the subtest set. For the feature selection step, we first performed a Pearson correlation between each feature and the SPE score for the samples in the subtraining set. A commonly used threshold ($p < 0.01$) was then applied to remove noisy edges and retain those that were significantly correlated (Rosenberg et al., 2016). These selected features were fed into the linear SVR models built with different values of C. Finally, according to the performance of these models on the subtest set, the best hyperparameter C was obtained and passed to the training set of the outer loop.

The outer loop adopted LOOCV to exploit every sample and evaluate the prediction performance more robustly (Vabalas et al., 2019). In each of the N iterations, a different participant was left out as the test set and the remaining N-1 as the training set. The training set then reselected features in a manner similar to the inner loop (Pearson correlation threshold $p < 0.01$). The retained edges were fed into the tuned linear SVR model with the best hyperparameter C determined by the inner loop. The predicted value for the test set was obtained by feeding the same features into the trained model. After N iterations, we acquired a predicted SPE score for all participants.

2.7. Prediction performance evaluation

Model performance (i.e., the correspondence between predicted and true values) was evaluated using Pearson correlations. The stronger the correlation between the predicted and true SPE scores, the more successful the model (Rosenberg et al., 2016). Notably, a common pitfall in machine learning is that each cross-validation fold does not satisfy independence, and the best practice for assessing the significance of model performance is permutation testing (Scheinost et al., 2019). Therefore, we randomly shuffled the phenotype (i.e., SPE score) and reran the prediction pipeline 1000 times to obtain a null distribution of the model performance. The *p*-value was calculated as (1 + the number of r-values in the null distribution that are greater than or equal to the r-value resulting from the unshuffled data)/1001 (Phipson and Smyth, 2010).

2.8. Functional anatomy

As each iteration of LOOCV may select slightly different features, we first identified those edges that appeared in all iterations to construct a consensus network for interpretation (see also Rosenberg et al., 2016). To quantify the contribution of each edge, linear machine-learning algorithms traditionally rank feature importance based on the raw weights assigned to features (Chang and Lin, 2008). However, Haufe et al. (2014) recently suggested that such an interpretation of the backward model (e.g., SVR and LASSO) may yield misleading conclusions. As a remedy for this case, they proposed the following formula to transform the weight matrix into an interpretable activation pattern matrix:

$$A = \sum_X W \sum_y^{-1},$$

where $\boldsymbol{A}$ represents the derived activation pattern matrix, $\boldsymbol{W}$ is the raw weight matrix of the linear SVR, $\sum_{\mathbf{X}}$ is the covariance matrix of the data, and $\sum_{\mathbf{y}}^{-1}$ denotes the inverse covariance of the prediction values.

Each edge in the consensus network's contribution was computed as the averaged activation pattern over iterations. We then summarized the connectivity patterns using two methods. First, we grouped the 246 nodes into 24 macroscale brain regions anatomically defined by the Brainnetome Atlas (Fan et al., 2016). The contribution of connectivity in each pair of macroscale regions was characterized as the sum of the activation patterns of all edges within it (for a similar approach, see Fong et al., 2019). Second, we regrouped the nodes into canonical networks based on the Yeo 17-network parcellation (Yeo et al., 2011). The mapping relationship between nodes and networks was defined by the atlas development team (Luo et al., 2020; Q. Wang et al., 2021). For clarity, we merged related sub-networks in the 17 networks (see also Shine et al., 2017) and referred to the undefined subcortical regions as subcortical networks. For the resulting nine canonical networks, we evaluated the contribution of within- or between-network connectivity analogously to the analyses of macroscale brain regions.

Moreover, to assess the role of each canonical network individually, we performed further computational lesion analysis. A lesion in a canonical network means that the signals of all its constituent nodes are erased, and the pipeline is then reran based on the resulting FC matrix (Rosenberg et al., 2016). The predictive performances of these "lesioned" brains were subsequently compared separately with the whole-brain result using Steiger's Z test.

2.9. Model robustness and specificity analyses

Previous studies have suggested that the predictive power of models may be affected by confounding factors (for an overview, see Scheinost et al., 2019). In the worst case, the model may predict not the desired phenotype but confounders such as demographic variables or head movement. To verify the robustness of the identified network to

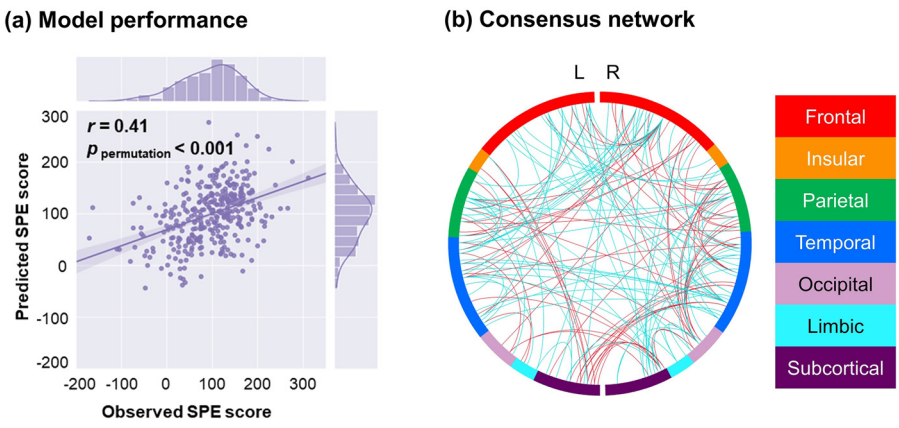

Fig. 2. Visualization of prediction results. (a) Scatter plot depicting the correlation between predicted and observed SPE score. (b) Circle plot showing the identified consensus network. The lines represent the connections between two nodes. Positive connections (i.e., edges significantly positively correlated with phenotype) are red, whereas negative connections are cyan. SPE, self-prioritization effect; L, left; R, right.

these factors, we performed several control analyses. First, instead of using Pearson's correlation, we selected the most relevant features based on partial correlation coefficients controlling for age, gender, or head movement (defined as mean framewise displacement). Second, before feature selection, we removed any edges associated with age or head movement or that differed between genders ($p < 0.01$). Third, we reconstructed a model using only these covariates as features and compared its performance with that of the main model.

Another consideration for robustness is the choice of the algorithm. Although SVR is one of the most widely used algorithms (Cui and Gong, 2018), we replaced it with several other common linear models to further boost robustness, including the ordinary least squares (OLS) regression, least absolute shrinkage and selection operator (LASSO), and ridge regression. Note that feature selection and CV strategies were the same.

Finally, to examine the specificity of the prediction model, we used the identified consensus network to predict participants' self-construal (Markus and Kitayama, 1991), which was measured using the 30-item Self-Construal Scale (Singelis, 1994). As self-construal is slightly associated with SPE (Jiang and Sui, 2022), unsatisfactory predictive performances can be expected if the consensus network is specific to the SPE score (Goldfarb et al., 2020).

2.10. External validation

To further assess the out-of-sample generalizability of the identified SPE-specific consensus network, we performed external validation analyses using an independent dataset. A detailed description of this dataset is available in the online supplementary materials. In brief, 66 healthy older adults were recruited to complete a 5-minute rs-fMRI scan and a 42-minute task-fMRI scan, during which they completed the perceptual matching task. We use the rs-fMRI data to predict SPE scores (defined identically to the main dataset) computed from behavioral data of task-fMRI. Since previous studies demonstrated that older adults retain similar SPE to the young (Gutchess et al., 2007; Mattan et al., 2017), we expected that the neural model from the main results could generalize to older adults to some extent.

The validation procedure was consistent with previous machine learning studies in neuroscience (e.g., Liu et al., 2021). We first excluded participants that did not participate in rs-fMRI scanning ($n = 2$) or were missing fMRI data in one or more nodes of the atlas ($n = 2$), resulting in 62 participants for validation (34 women; mean age = 71.3 years; age range = 58–84 years). Their rs-fMRI data were preprocessed using the same pipeline as the main dataset to obtain FC matrices. Subsequently, we refit the main dataset (all 348 participants) using the identified consensus network as features to learn the parameters of the model. The independent dataset's FC matrices were then fed into the trained model to generate predicted SPE scores. Considering that using an older sample may introduce additional confounds, we further evaluated validation performance using partial correlation, controlling for age, head movement, and gender. Notably, since older adults generally move their heads more than younger adults, using the same excluding criterion as the main dataset (i.e., > 0.2 mm mean framewise displacement) would eliminate nearly half of the samples ($n = 27$). Therefore, we did not exclude older subjects with excessive head movement to preserve statistical power. However, the results under this exclusion criterion were also reported.

3. Results

3.1. Predictability of SPE

As shown in Fig. 2a, in the unseen data (i.e., not used to train the model), the true value of the SPE was significantly correlated with the predicted value ($r = 0.41$, permutation $p < 0.001$, MAE = 52.80), indicating that our linear SVR model successfully predicted the SPE of individuals.

3.2. Contributing networks to prediction

In the nested LOOCV, the number of selected features varied from 317 to 383 in multiple iterations, representing less than 2% of the brain's 30,135 total edges as defined by the atlas. The identified consensus network that contained 314 edges was visualized in Fig. 2b. Consistent with previous prediction studies (Yip et al., 2019; Zhang et al., 2018), the highest-degree nodes (i.e., nodes with the most connections) in the consensus network were widely distributed across the brain, including the frontal, occipital, parietal, and temporal lobes, as well as the subcortical regions (see Table S1). Most of these high-degree nodes were located in the right hemisphere. In addition, we listed the top 20 positive and negative FCs with the strongest predictive contribution in Tables S2 and S3, respectively. As shown in the tables, FCs related to self-processing tended to involve regions such as the medial prefrontal cortex (mPFC) or subcortical structures, while FCs related to other-processing tended to involve regions such as the lateral PFC or posterior cingulate cortices.

The contribution (i.e., summarized activation pattern) of FCs among the 24 macroscale brain regions was depicted in Fig. 3. The results indicated that the FCs between the thalamus and the parahippocampal gyrus (PhG), lateral occipital cortex (LOcC) and superior temporal gyrus (STG), thalamus and insular gyrus, superior frontal gyrus (SFG) and inferior frontal gyrus (IFG) were the primary predictors of stronger self-prioritization. Meanwhile, the FCs between the PhG and middle temporal gyrus, PhG and middle frontal gyrus, and within the IFG predicted weaker self-prioritization.

Besides, we also regrouped the 264 nodes into nine canonical networks. The contribution of FCs among those canonical networks was

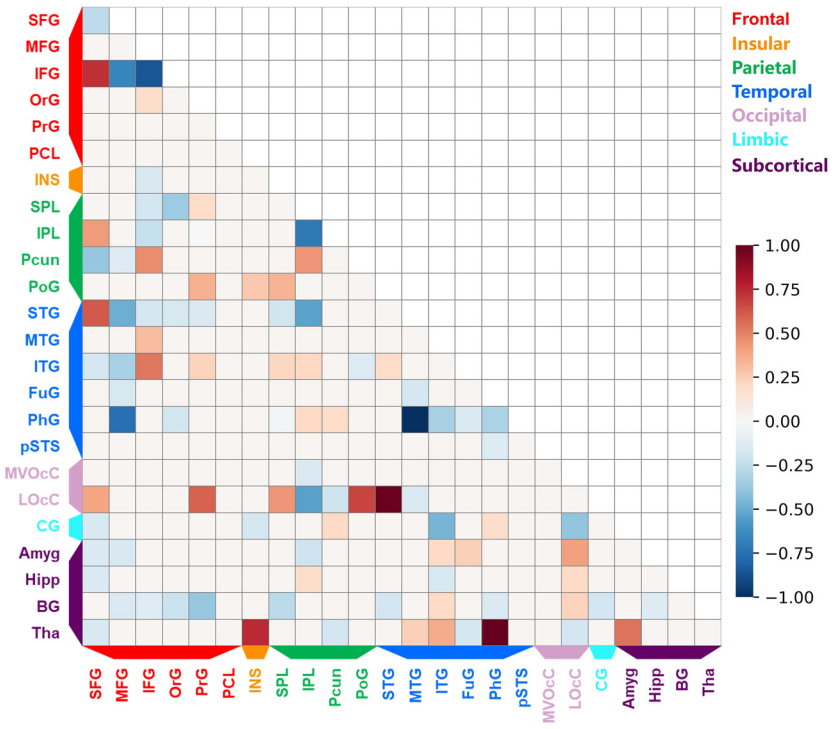

Fig. 3. The contributions of FCs between each pair of macroscale brain regions to the predictive model. Cells on the diagonal represent the results of FCs within macroscale brain regions. The values in each cell denote the sum of activation patterns. For better readability, these values are normalized to between −1 and 1 using scikit-learn's MinMaxScaler module. SFG, superior frontal gyrus; MFG, middle frontal gyrus; IFG, inferior frontal gyrus; OrG, orbital gyrus; PrG, precentral gyrus; PCL, paracentral lobule; INS, insular gyrus; SPL, superior parietal lobule; IPL, inferior parietal lobule; Pcun, precuneus; PoG, postcentral gyrus; STG, superior temporal gyrus; MTG, middle temporal gyrus; ITG, inferior temporal gyrus; FuG, fusiform gyrus; PhG, parahippocampal gyrus; pSTS, posterior superior temporal sulcus; MVOcC, medioventral occipital cortex; LOcC, lateral occipital cortex; CG, cingulate gyrus; Amyg, amygdala; Hipp, hippocampus; BG, basal ganglia; Tha, thalamus.

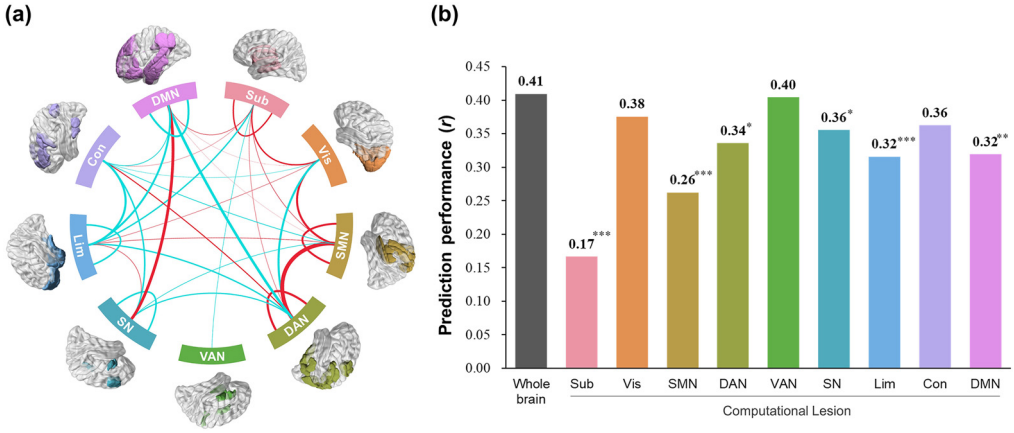

Fig. 4. Network-level results. (a) The sum of the FC's activation pattern for each pair of canonical networks. Positive connections are red, whereas negative connections are cyan. Thicker and darker connections indicate stronger connection strength (i.e., higher absolute value of the activation pattern). (b) Model performance of whole-brain FC and after computational lesion of each canonical network. The significance symbols indicate whether there is a significant difference between each computational lesioned model and the whole-brain model. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Sub, subcortical network; Vis, visual network; SMN, somatomotor network; DAN, dorsal attention network; VAN, ventral attention network; SN, salience network; Lim, limbic network; Con, control network; DMN, default mode network.

visualized in Fig. 4a. The results demonstrated the predictive power of connectivity between the dorsal attention network (DAN) and somatomotor network (SMN), the DMN and salience network (SN), and the DMN and DAN. As for the individual importance of each canonical network, the computational lesion analysis results (see Fig. 4b) revealed that the model could still significantly predict individuals' SPE scores without any single canonical network, which corresponds to findings in other fields, confirming that machine learning models tend to utilize whole-brain signals (Feng et al., 2018). Nevertheless, Steiger's Z test still found several canonical networks that caused significant degradation in model performance after computational lesions, including the subcortical network ($z = 5.74$, $p < 0.001$), limbic network ($z = 3.30$, $p < 0.001$), SMN ($z = 4.15$, $p < 0.001$), DMN ($z = 2.85$, $p < 0.01$), DAN ($z = 2.2$, $p = 0.03$), and SN ($z = 2.35$, $p = 0.02$).

3.3. Model robustness, specificity, and generalizability

As described in the Methods section, we performed several control analyses to test the robustness of the identified network to confounding factors. First, when selecting features based on partial correlation coefficients that controlled for age, gender, or head movement, the resulting networks still significantly predicted the participants' SPE scores ($rs > 0.25$, permutation $ps < 0.001$). Second, although we removed any edges associated with age or head movement or that differed between genders, the streamlined networks still significantly predicted the SPE ($rs > 0.31$, permutation $ps < 0.001$). Third, the covariate-only model could not predict participants' SPE scores ($r = -0.16$, permutation $p = 0.99$), and Steiger's Z test showed that its performance was significantly worse than that of the main model ($z = 7.78$, $p < 0.001$).

The results from different algorithms revealed that the predictions of the LASSO and ridge regression remained significant after replacing the SVR ($r = 0.21$ and 0.23, respectively, permutation $ps < 0.001$), whereas the OLS regression failed to predict individuals' SPE scores ($r = 0.08$, permutation $p = 0.07$). The inferior performance of LASSO and ridge regression compared to SVR echoes previous machine-learning research (Cho et al., 2020). The unsatisfactory performance of OLS can be attributed to overfitting caused by a lack of regularization (Ying, 2019).

Regarding the issue of GSR, in line with previous machine-learning based studies (e.g., Ju et al., 2020), the performance of the SVR model using FC matrices without GSR was not significant ($r = -0.05$, permutation $p = 0.82$). As suggested by Murphy and Fox (2017), GSR is not inherently right or wrong and depends on research questions and methods. A recent large-scale investigation by Li et al. (2019) found that GSR sharply strengthens the relationship between rs-FC and behavioral measures, which provides explanations for our results and the general use of GSR in machine-learning research (Goldfarb et al., 2020; Yu et al., 2020).

The results of the specificity test demonstrated that although the consensus network could successfully predict SPE scores, it could not significantly predict participants' independent self ($r = -0.16$, permutation $p > 0.99$) or interdependent self ($r = 0.04$, permutation $p = 0.14$). These results demonstrate that the consensus network is specific to the SPE, which is a behavior-level self-bias and cannot be generalized to higher-level self-reflections such as self-construals. Interestingly, a previous study demonstrated that self-interdependency has a stronger impact on individuals' self-prioritization than self-independence (Jiang and Sui, 2022). This is consistent with our results, showing that the interdependent self predicts slightly better than the independent self ($z = 2.70$, $p < 0.01$).

Finally, external validation results indicate that the SPE-specific network and model identified in the main dataset ($n = 348$) can be generalized to an independent dataset of older adults [$r(62) = 0.33$, permutation $p < 0.01$, MAE = 217.2, see Fig. S2]. To ensure that confounding variables did not play a role, we calculated the partial correlation between predicted and actual SPE scores controlling for age, gender, and head movement. The predictive performance showed little change [$r(62) = 0.32$, permutation $p < 0.01$]. In addition, after removing participants with excessive head movement (using the same criteria as the main dataset), the predictive performance remained significant [$r(35) = 0.30$, permutation $p = 0.049$, MAE = 201.3].

4. Discussion

Longstanding controversies and disparate theories regarding the emergence of the sense of self have necessitated more neuroscientific evidence. Here, we demonstrated how behavioral results in self-prioritization could be predicted via correlations with spontaneous brain resting-state activity, indicating readiness for self-related processes in the brain at rest. The results demonstrated that the trained model was robust and that the identified predictive networks were SPE-specific. The predictions were not confounded by covariates such as head movements and were consistent across different algorithms. Most importantly, the generalization ability of the predictive model was externally validated in an independent dataset. Our results echo the expected manifestations of the fundamental self hypothesis, thus providing preliminary evidence for the fundamental self hypothesis.

This study has significant theoretical implications in two aspects. First, we have provided direct evidence showing that self-specificity is indeed embedded within the brain's spontaneous brain activities, supporting the BMSS's rest-self containment assumption instead of the rest-self overlap assumption. Second, we have used unbiased data-driven technologies to demonstrate the association between self-specificity and subcortical structures, which has not been thoroughly investigated previously. This, however, raises two inevitable questions regarding the fundamental self hypothesis. On the one hand, how 'fundamental' can this self component be? In extreme cases, as Northoff and Panksepp (2008) suggested, the fundamental component of self might be possible to exist in species other than humans as humans and primates share similar subcortical structures (Northoff and Panksepp, 2008) and RSNs (Rilling et al., 2007). On the other hand, how does the fundamental self associate with perhaps one of the highest-order brain functions, namely self-consciousness? Given the complexity of the self, it should be cautious when assuming that all self components operate at the fundamental connection level.

Focusing on the self-processing aspects of the fundamental self, in the following subsections, we will discuss the neural characteristics revealed in the current study and develop a novel neural model to explain the fundamental self hypothesis.

4.1. Brain region patterns of SPE-specific network

A closer look at the consensus predictive network revealed that the FCs of several nodes within the DMN, SN, and DAN over the frontal and temporal lobes, such as the SFG (including nodes of the bilateral ventromedial prefrontal cortex [vmPFC]), IFG, and STG, contributed significantly to the model. The involvement of these regions in self-related cognition has been repeatedly reported (for meta-analyses, see Hu et al., 2016; Murray et al., 2012). Our results confirmed these regions' predictive role in self-processing by echoing studies using multivoxel pattern analysis (MVPA) techniques to decode differences between self and others (Woźniak et al., 2022; Yankouskaya et al., 2017). In addition to these high-level cortices, the current findings highlight the role of several subcortical structures or adjacent areas, especially the FCs between the thalamus, insula, amygdala, and PhG. Although not traditionally recognized as neural substrates of the self, these regions have been frequently observed in empirical self-processing studies (Murray et al., 2015; Qin et al., 2020; Sui and Gu, 2017). Recently, Qin et al. (2020) conducted a meta-analysis to test a three-level-self model elucidating how the bodily interoceptive self is integrated within the self, highlighting the role of these subcortical regions in interoceptive-self processing. Similarly, this integrative view was supported by another line of empirical work by Babo-Rebelo and colleagues (Babo-Rebelo et al., 2019; Babo-Rebelo et al., 2016a, 2016b), which revealed functional couplings between self-processing (via higher-order cortices) and visceral signal monitoring (via subcortical structures).

However, the precise contributions of these subcortical regions can vary from one another. For example, the insula may couple the neural monitoring of cardiac signals to the self (Babo-Rebelo et al., 2016), whereas the amygdala's involvement may stem from individuals' affective-manner processing of self-related stimuli (Rameson et al., 2010). The noteworthy negative connections of the PhG may be due to its episodic memory function (Hayes et al., 2007). That is, similar to attentional network activation (Humphreys and Sui, 2016), the brain may require more intense PhG activation in encoding and retrieving episodic memories related to others.

The thalamus, the subcortical structure with the most noticeable results, is sometimes noted (Murray et al., 2015; Northoff et al., 2009), but its function remains largely unexplained. The thalamus has traditionally been described as a machine-like relay station for cortical information (Sherman, 2007). However, contemporary perspectives suggest that it may continue to play an essential role in subsequent sensory and motor cortical processing through transthalamic corticocortical pathways (Sherman, 2016; Worden et al., 2021). Shine (2021) even argued that the thalamus might be responsible for integrating and regulating the macrosystems of the brain. Correspondingly, we proposed that the thalamus may continuously be involved in the self's readiness to regulate cognitive processing.

Unexpectedly, the LOcC also showed closed connections to other regions. The involvement of visual cortices in self-related processing has previously been reported (Ferri et al., 2012; Hu et al., 2016; Murray et al., 2015). Using MVPA techniques, Yin et al. (2021) re-

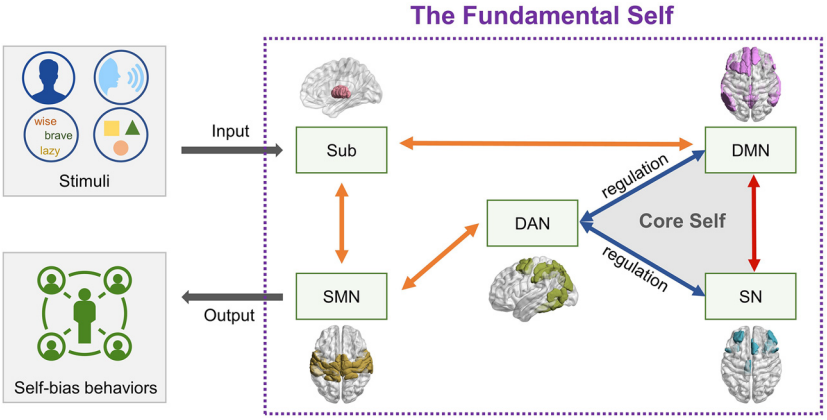

Fig. 5. The intrinsic self-processing network model. The fundamental self is assumed to be topologically represented by the interactions among these five networks, three of which constitute the traditionally recognized core self circuit (i.e., DMN-SN-DAN). Visual or auditory stimuli are first transmitted through the thalamus to higher-order cortices. We speculate that the SMN serves as a bridge between the mental and bodily selves and is involved in generating observable self-biased behaviors. Sub, subcortical network; SMN, somatomotor network; DAN, dorsal attention network; DMN, default mode network; SN, salience network.

cently suggested that visual cortices embedded the degree of prioritization of self- or other-associated stimuli in working memory. Another possibility for the widespread involvement of the LOcC may be attributed to its modulating role in the visual processing of shapes (Cant and Goodale, 2007). However, despite these connections, the computational lesion of the visual network did not significantly affect the model's performance. We speculate that visual functions may not be part of the fundamental self and thus serve as a fine but unnecessary factor in SPE prediction. This view is consistent with a previous study showing that patients with visual extinction still had an SPE (Sui and Humphreys, 2017a).

4.2. Functional system patterns and a novel neural model of the self

When organizing the brain into large-scale functional systems, the contributing FCs are widely distributed among the brain's canonical networks. This was echoed by computational lesion analysis, which showed that six of the nine networks contributed significantly to the model's predictive performance. This widespread pattern of results has two implications. First, although the triad of DMN-SN-DAN has been addressed by previous theoretical models that account for the neural basis of the self (Northoff, 2016; Sui and Gu, 2017), the contributions of the limbic network, SMN, and subcortical network have been less frequently discussed. In our computational lesion analysis, the impact of the SMN and subcortical network on model performance was much greater than that of the DMN, SN, and DAN, indicating that they also play significant roles in the self-prioritization process.

To synthesize previous theoretical works (Northoff, 2016; Qin et al., 2020; Sui and Gu, 2017), as well as the current findings, we propose an intrinsic self-processing network model (see Fig. 5) that describes the neural mechanism that gives rise to the self as object (Sui and Gu, 2017) and regulates other cognitive functions. Based on the BMSS model (Northoff, 2016), it is believed that self-specificity is embedded in the brain's resting-state spontaneous activity. Corroborating this viewpoint, the findings have demonstrated that this intrinsic self-processing network involves multiple neural couplings between three neural networks and communications between these networks and sub-cortical structures. In particular, self-specificity is associated with the intrinsic positive couplings between DMN and SN and the intrinsic negative couplings between DAN and DMN/SN. These cortical network connections are consistent with the self-as-object model proposing that the self emerges from the interaction between the 'self' network (e.g., vmPFC), cognitive control network (e.g., dorsolateral PFC), and salience network (e.g., insula) (Sui and Gu, 2017). On the other hand, the result in terms of communications between cortical networks, subcortical network, and SMN is consistent with the three-level-self view (Qin et al., 2020) that observable self-biased behaviors employ mental self-processing and rely on endowing non-bodily external stimulus self-specificity and extensive brain involvement. The role of the SMN may reflect that the bodily self also derives from somatomotor experiences (Ferri et al., 2012) and serves as an essential component of self functions (Legrand, 2006).

4.3. Implications of the intrinsic self-processing network model

The intrinsic self-processing network model can be used to interpret previous findings. On the one hand, according to this model, individuals with greater self-specificity in the brain's spontaneous activity may have stronger connectivity between the DMN and SN. The DMN is often linked to self-related processing owing to the anatomical overlap of brain regions (Davey et al., 2016; Northoff et al., 2006; Qin and Northoff, 2011). Similarly, representative regions of the SN, such as the insula and anterior cingulate, are consistently involved in self-related processing (Qin et al., 2020; Scalabrini et al., 2021). The broad involvement of the SN can be attributed to the fact that the self is a salient stimulus (Humphreys and Sui, 2015; Tacikowski and Nowicka, 2010). However, although DMN-SN interaction has been linked to many brain functions (Doll et al., 2015; Jilka et al., 2014), it has only recently been used to characterize the self (Sui and Gu, 2017). In line with their hypothesis and our model, Yankouskaya and Sui (2022) observed stronger DMN-SN connectivity activity during self-processing than other-processing.

On the other hand, our model indicates that individuals with stronger FC between the DAN and DMN or SN possess a weaker self-prioritization tendency. Consistently, Sui et al. (2013) found that the dorsolateral PFC, a key node of the DAN, showed enhanced activation in other-reference processing compared to self-related processing. Later, Sui et al. (2015) observed a hyperself-bias effect on the attentional control network in a patient with lesions. They then proposed a self-attention network to explain the role of attention in social processing (Humphreys and Sui, 2016). Based on evidence from multiple domains, they concluded that self-related stimuli are naturally salient and attention-attracting; therefore, individuals need to mobilize a top-down attentional system to suppress spontaneous self-bias when switching to other-related stimuli. This echoes our model that weaker self-bias may also stem from the stronger inhibitory capacity of the attention network.

4.4. Limitations and future directions

In summary, this study takes the first step in using a data-driven machine-learning approach to predict individuals' self-prioritization, supporting the hypothesis that self-specific information is embedded within the spontaneous activity of the brain (Northoff, 2016). Although it contributes to our knowledge of the neural substrate of the self, the present study has several limitations that need to be acknowledged. First, although our sample size was relatively large in neuroscience research, the prediction stability of machine learning improves significantly as sample size increases (Jollans et al., 2019). Given the high

cross-site similarity of rs-fMRI, the predictive power of the identified networks can be validated using more diverse samples. Second, we used a perceptual matching task to measure self-bias because it excluded the effect caused by differences in stimulus familiarity (Woźniak and Knoblich, 2019). One promising direction is to have participants complete various self/other processing trait-like tasks and then refine the core self-representing brain regions by constructing multiple predictive models. Another trend in machine learning-based neuroimaging research is the use of multi-modal features for prediction, such as task-fMRI (Rosenberg et al., 2016), dynamic FC (Fong et al., 2019), and structural connectivity (Yu et al., 2020). Future studies could incorporate multi-modal data into the model to determine whether better biomarkers can be identified.

Data and code availability statement

Preprocessed image data, behavioral data, and analysis scripts have been made publicly available via OSF and can be accessed at https://osf.io/hbrus/?view_only=98b2095f72e64fd382fc0d33de3f4497. The Human Brainnetome Atlas is available online at https://atlas.brainnetome.org/download.html.

Declaration of Competing Interest

The authors declare no conflicts of interest.

Credit authorship contribution statement

Yongfa Zhang: Methodology, Software, Formal analysis, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Fei Wang:** Conceptualization, Methodology, Investigation, Visualization, Writing – review & editing, Supervision, Project administration. **Jie Sui:** Conceptualization, Writing – review & editing, Supervision, Project administration.

Acknowledgments

Fei Wang acknowledges the financial support from The Research Project of Shanghai Science and Technology Commission (20dz2260300). Jie Sui acknowledges the financial support from the Leverhulme Trust (RPG-2019–010).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.120205.

References

Babo-Rebelo, M., Buot, A., Tallon-Baudry, C., 2019. Neural responses to heartbeats distinguish self from other during imagination. Neuroimage 191, 10–20.
Babo-Rebelo, M., Richter, C.G., Tallon-Baudry, C., 2016a. Neural responses to heartbeats in the default network encode the self in spontaneous thoughts. J. Neurosci. 36 (30), 7829–7840.
Babo-Rebelo, M., Wolpert, N., Adam, C., Hasboun, D., Tallon-Baudry, C., 2016b. Is the cardiac monitoring function related to the self in both the default network and right anterior insula? Philos. Trans. Royal Society B: Biol. Sci. 371 (1708), 20160004.
Cant, J.S., Goodale, M.A., 2007. Attention to form or surface properties modulates different regions of human occipitotemporal cortex. Cereb. Cortex 17 (3), 713–731. doi:10.1093/cercor/bhk022.
Chang, Y., Lin, C., 2008. Feature Ranking Using Linear SVM. In: Proceedings of the Workshop on the Causation and Prediction Challenge at WCCI 2008, 3, pp. 53–64.
Cho, B.H., Koyama, K., Diaz, E.O., Koseki, S., 2020. Determination of "hass" avocado ripeness during storage based on smartphone image and machine learning model. Food Bioproc. Tech. 13 (9), 1579–1587. doi:10.1007/s11947-020-02494-x.
Cui, Z., Gong, G., 2018. The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. Neuroimage 178, 622–637.
Cunningham, S.J., Turk, D.J., 2017. A review of self-processing biases in cognition. Q. J. Exp. Psychol. 70 (6), 987–995.
Damasio, A.R., 1999. How the brain creates the mind. Sci. Am. 281 (6), 112–117.
Davey, C.G., Pujol, J., Harrison, B.J., 2016. Mapping the self in the brain's default mode network. Neuroimage 132, 390–397.
de Greck, M., Rotte, M., Paus, R., Moritz, D., Thiemann, R., Proesch, U., Bruer, U., Moerth, S., Tempelmann, C., Bogerts, B., 2008. Is our self based on reward? Self-relatedness recruits neural activity in the reward system. Neuroimage 39 (4), 2066–2075.
Doll, A., Hölzel, B.K., Boucard, C.C., Wohlschläger, A.M., Sorg, C., 2015. Mindfulness is associated with intrinsic functional connectivity between default mode and salience networks. Front. Hum. Neurosci. 9, 461.
Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A.R., 2016. The human brainnetome atlas: a new brain atlas based on connectional architecture. Cereb. Cortex 26 (8), 3508–3526.
Feng, C., Yuan, J., Geng, H., Gu, R., Zhou, H., Wu, X., Luo, Y., 2018. Individualized prediction of trait narcissism from whole-brain resting-state functional connectivity. Hum. Brain Mapp. 39 (9), 3701–3712.
Ferri, F., Frassinetti, F., Ardizzi, M., Costantini, M., Gallese, V., 2012. A sensorimotor network for the bodily self. J. Cogn. Neurosci. 24 (7), 1584–1595.
Fong, A.H.C., Yoo, K., Rosenberg, M.D., Zhang, S., Li, C.-S.R., Scheinost, D., Constable, R.T., Chun, M.M., 2019. Dynamic functional connectivity during task performance and rest predicts individual differences in attention across studies. Neuroimage 188, 14–25.
Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., Turner, R., 1996. Movement-related effects in fMRI time-series. Magn. Reson. Med. 35 (3), 346–355.
Gallagher, S., Meltzoff, A.N., 1996. The earliest sense of self and others: merleau-Ponty and recent developmental studies. Philos. Psychol. 9 (2), 211–233.
Goldfarb, E.V., Rosenberg, M.D., Seo, D., Constable, R.T., Sinha, R., 2020. Hippocampal seed connectome-based modeling predicts the feeling of stress. Nat. Commun. 11 (1), 1–10.
Gutchess, A.H., Kensinger, E.A., Schacter, D.L., 2007. Aging, self-referencing, and medial prefrontal cortex. Soc. Neurosci. 2 (2), 117–133.
Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., Bießmann, F., 2014. On the interpretation of weight vectors of linear models in multivariate neuroimaging. Neuroimage 87, 96–110.
Hayes, S.M., Nadel, L., Ryan, L., 2007. The effect of scene context on episodic object recognition: parahippocampal cortex mediates memory encoding and retrieval success. Hippocampus 17 (9), 873–889. doi:10.1002/hipo.20319.
Hu, C., Di, X., Eickhoff, S.B., Zhang, M., Peng, K., Guo, H., Sui, J., 2016. Distinct and common aspects of physical and psychological self-representation in the brain: a meta–analysis of self-bias in facial and self-referential judgements. Neurosci. Biobehav. Rev. 61, 197–207.
Huang, Z., Obara, N., Davis IV, H.H., Pokorny, J., Northoff, G, 2016. The temporal structure of resting-state brain activity in the medial prefrontal cortex predicts self-consciousness. Neuropsychologia 82, 161–170.
Humphreys, G.W., Sui, J., 2015. The salient self: social saliency effects based on self-bias. J. Cognitive Psychol. 27 (2), 129–140.
Humphreys, G.W., Sui, J., 2016. Attentional control and the self: the Self-Attention Network (SAN). Cogn. Neurosci. 7 (1–4), 5–17.
Jiang, M., Sui, J., 2022. Bicultural minds: a cultural priming approach to the self-bias effect. Behav. Sci. 12 (2), 45.
Jiang, R., Calhoun, V.D., Zuo, N., Lin, D., Li, J., Fan, L., Qi, S., Sun, H., Fu, Z., Song, M., 2018. Connectome-based individualized prediction of temperament trait scores. Neuroimage 183, 366–374.
Jilka, S.R., Scott, G., Ham, T., Pickering, A., Bonnelle, V., Braga, R.M., Leech, R., Sharp, D.J., 2014. Damage to the salience network and interactions with the default mode network. J. Neurosci. 34 (33), 10798–10807.
Jollans, L., Boyle, R., Artiges, E., Banaschewski, T., Desrivières, S., Grigis, A., Martinot, J.-L., Paus, T., Smolka, M.N., Walter, H., 2019. Quantifying performance of machine learning methods for neuroimaging data. Neuroimage 199, 351–365.
Ju, Y., Horien, C., Chen, W., Guo, W., Lu, X., Sun, J., Dong, Q., Liu, B., Liu, J., Yan, D., 2020. Connectome-based models can predict early symptom improvement in major depressive disorder. J. Affect. Disord. 273, 442–452.
Kircher, T.T., Senior, C., Phillips, M.L., Benson, P.J., Bullmore, E.T., Brammer, M., Simmons, A., Williams, S.C., Bartels, M., David, A.S., 2000. Towards a functional neuroanatomy of self processing: effects of faces and words. Cognitive Brain Res. 10 (1–2), 133–144.
Kotlewska, I., Nowicka, A., 2015. Present self, past self and close-other: event-related potential study of face and name detection. Biol. Psychol. 110, 201–211.
Kotlewska, I., Nowicka, A., 2016. Present-self, past-self and the close-other: neural correlates of assigning trait adjectives to oneself and others. Eur. J. Neurosci. 44 (4), 2064–2071.
Larabi, D.I., Renken, R.J., Cabral, J., Marsman, J.-B.C., Aleman, A., Ćurčić-Blake, B., 2020. Trait self-reflectiveness relates to time-varying dynamics of resting state functional connectivity and underlying structural connectomes: role of the default mode network. Neuroimage 219, 116896.
Legrand, D., 2006. The bodily self: the sensori-motor roots of pre-reflective self-consciousness. Phenomenol. Cognitive Sci. 5 (1), 89–118.
Li, J., Kong, R., Liégeois, R., Orban, C., Tan, Y., Sun, N., Holmes, A.J., Sabuncu, M.R., Ge, T., Yeo, B.T., 2019. Global signal regression strengthens association between resting-state functional connectivity and behavior. Neuroimage 196, 126–141.
Liu, P., Yang, W., Zhuang, K., Wei, D., Yu, R., Huang, X., Qiu, J., 2021. The functional connectome predicts feeling of stress on regular days and during the COVID-19 pandemic. Neurobiol. Stress 14, 100285.
Luo, N., Sui, J., Abrol, A., Chen, J., Turner, J.A., Damaraju, E., Fu, Z., Fan, L., Lin, D., Zhuo, C., 2020. Structural brain architectures match intrinsic functional networks and vary across domains: a study from 15 000+ individuals. Cerebral Cortex 30 (10), 5460–5470.

Macrae, C.N., Visokomogilski, A., Golubickis, M., Sahraie, A., 2018. Self-relevance enhances the benefits of attention on perception. Vis. cogn. 26 (7), 475–481.
Magnusson, M., Andersen, M., Jonasson, J., Vehtari, A., 2019. Bayesian leave-one-out cross-validation for large data. In: Proceedings of the 36th International Conference on Machine Learning, 97, pp. 4244–4253.
Markus, H.R., Kitayama, S., 1991. Culture and the self: implications for cognition, emotion, and motivation. Psychol. Rev. 98 (2), 224.
Mattan, B.D., Quinn, K.A., Acaster, S.L., Jennings, R.M., Rotshtein, P., 2017. Prioritization of self-relevant perspectives in ageing. Q. J. Experiment. Psychol. 70 (6), 1033–1052.
Moray, N., 1959. Attention in dichotic listening: affective cues and the influence of instructions. Q. J. Experiment. Psychol. 11 (1), 56–60.
Murphy, K., Fox, M.D., 2017. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. Neuroimage 154, 169–173.
Murray, R.J., Debbane, M., Fox, P.T., Bzdok, D., Eickhoff, S.B., 2015. Functional connectivity mapping of regions associated with self-and other-processing. Hum. Brain Mapp. 36 (4), 1304–1324.
Murray, R.J., Schaer, M., Debbané, M., 2012. Degrees of separation: a quantitative neuroimaging meta-analysis investigating self-specificity and shared neural activation between self-and other-reflection. Neurosci. Biobehav. Rev. 36 (3), 1043–1059.
Northoff, G., 2012. Immanuel Kant's mind and the brain's resting state. Trends Cogn. Sci. (Regul. Ed.) 16 (7), 356–359.
Northoff, G., 2016. Is the self a higher-order or fundamental function of the brain? The "basis model of self-specificity" and its encoding by the brain's spontaneous activity. Cogn. Neurosci. 7 (1–4), 203–222.
Northoff, G., Bermpohl, F., 2004. Cortical midline structures and the self. Trends Cogn. Sci. (Regul. Ed.) 8 (3), 102–107.
Northoff, G., Heinzel, A., De Greck, M., Bermpohl, F., Dobrowolny, H., Panksepp, J., 2006. Self-referential processing in our brain—A meta-analysis of imaging studies on the self. Neuroimage 31 (1), 440–457.
Northoff, G., Panksepp, J., 2008. The trans-species concept of self and the subcortical—cortical midline system. Trends Cogn. Sci. (Regul. Ed.) 12 (7), 259–264.
Northoff, G., Schneider, F., Rotte, M., Matthiae, C., Tempelmann, C., Wiebking, C., Bermpohl, F., Heinzel, A., Danos, P., Heinze, H.J., 2009. Differential parametric modulation of self-relatedness and emotions in different brain regions. Hum. Brain Mapp. 30 (2), 369–382.
Northoff, G., Vatansever, D., Scalabrini, A., Stamatakis, E.A., 2022. Ongoing brain activity and its role in cognition: dual versus baseline models. The Neuroscientist, 10738584221081752.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: machine learning in Python. J. Mach. Learning Res. 12, 2825–2830.
Phipson, B., Smyth, G.K., 2010. Permutation P-values should never be zero: calculating exact p-values when permutations are randomly drawn. Stat. Appl. Genet. Mol. Biol. 9 (1). doi:10.2202/1544-6115.1585.
Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59 (3), 2142–2154.
Qin, P., Grimm, S., Duncan, N.W., Fan, Y., Huang, Z., Lane, T., Weng, X., Bajbouj, M., Northoff, G., 2016. Spontaneous activity in default-mode network predicts ascription of self-relatedness to stimuli. Soc. Cogn. Affect. Neurosci. 11 (4), 693–702.
Qin, P., Northoff, G., 2011. How is our self related to midline regions and the default-mode network? Neuroimage 57 (3), 1221–1233.
Qin, P., Wang, M., Northoff, G., 2020. Linking bodily, environmental and mental states in the self—A three-level model based on a meta-analysis. Neurosci. Biobehav. Rev. 115, 77–95.
Rameson, L.T., Satpute, A.B., Lieberman, M.D., 2010. The neural correlates of implicit and explicit self-relevant processing. Neuroimage 50 (2), 701–708.
Rilling, J.K., Barks, S.K., Parr, L.A., Preuss, T.M., Faber, T.L., Pagnoni, G., Bremner, J.D., Votaw, J.R., 2007. A comparison of resting-state brain activity in humans and chimpanzees. Proceed. Nat. Acad. Sci. 104 (43), 17146–17151.
Rochat, P., 2003. Five levels of self-awareness as they unfold early in life. Conscious. Cogn. 12 (4), 717–731.
Rosenberg, M.D., Finn, E.S., Scheinost, D., Papademetris, X., Shen, X., Constable, R.T., Chun, M.M., 2016. A neuromarker of sustained attention from whole-brain functional connectivity. Nat. Neurosci. 19 (1), 165–171.
Scalabrini, A., Vai, B., Poletti, S., Damiani, S., Mucci, C., Colombo, C., Zanardi, R., Benedetti, F., Northoff, G., 2020. All roads lead to the default-mode network—Global source of DMN abnormalities in major depressive disorder. Neuropsychopharmacology 45 (12), 2058–2069.
Scalabrini, A., Wolman, A., Northoff, G., 2021. The self and its right insula—Differential topography and dynamic of right vs. left insula. Brain Sci. 11 (10), 1312.
Scheinost, D., Noble, S., Horien, C., Greene, A.S., Lake, E.M., Salehi, M., Gao, S., Shen, X., O'Connor, D., Barron, D.S., 2019. Ten simple rules for predictive modeling of individual differences in neuroimaging. Neuroimage 193, 35–45.
Scheller, M., Sui, J., 2022. The power of the self: anchoring information processing across contexts. J. Experiment. Psychol.: Hum. Percept. Perform..
Schneider, F., Bermpohl, F., Heinzel, A., Rotte, M., Walter, M., Tempelmann, C., Wiebking, C., Dobrowolny, H., Heinze, H., Northoff, G., 2008. The resting brain and our self: self-relatedness modulates resting state neural activity in cortical midline structures. Neuroscience 157 (1), 120–131.
Sherman, S.M., 2007. The thalamus is more than just a relay. Curr. Opin. Neurobiol. 17 (4), 417–422.
Sherman, S.M., 2016. Thalamus plays a central role in ongoing cortical functioning. Nat. Neurosci. 19 (4), 533–541.
Shine, J.M., 2021. The thalamus integrates the macrosystems of the brain to facilitate complex, adaptive brain network dynamics. Prog. Neurobiol. 199, 101951.
Shine, J.M., Kucyi, A., Foster, B.L., Bickel, S., Wang, D., Liu, H., Poldrack, R.A., Hsieh, L.-T., Hsiang, J.C., Parvizi, J., 2017. Distinct patterns of temporal and directional connectivity among intrinsic networks in the human brain. J. Neurosci. 37 (40), 9667–9674.
Singelis, T.M., 1994. The measurement of independent and interdependent self-construals. Personal. Soc. Psychol. Bull. 20 (5), 580–591.
Sripada, C., Rutherford, S., Angstadt, M., Thompson, W.K., Luciana, M., Weigard, A., Hyde, L.H., Heitzeg, M., 2020. Prediction of neurocognition in youth from resting state fMRI. Mol. Psychiatry 25 (12), 3413–3421.
Stolte, M., Humphreys, G., Yankouskaya, A., Sui, J., 2017. Dissociating biases towards the self and positive emotion. Q. J. Experiment. Psychol. 70 (6), 1011–1022.
Sui, J., Enock, F., Ralph, J., Humphreys, G.W., 2015. Dissociating hyper and hypoself biases to a core self-representation. Cortex 70, 202–212.
Sui, J., Gu, X., 2017. Self as object: emerging trends in self research. Trends Neurosci. 40 (11), 643–653.
Sui, J., He, X., Humphreys, G.W., 2012. Perceptual effects of social salience: evidence from self-prioritization effects on perceptual matching. J. Experiment. Psychol.: Hum. Percept. Perform. 38 (5), 1105.
Sui, J., Humphreys, G.W., 2015. The integrative self: how self-reference integrates perception and memory. Trends Cogn. Sci. (Regul. Ed.) 19 (12), 719–728.
Sui, J., Humphreys, G.W., 2017a. The self survives extinction: self-association biases attention in patients with visual extinction. Cortex 95, 248–256.
Sui, J., Humphreys, G.W., 2017b. The ubiquitous self: what the properties of self-bias tell us about the self. Ann. N. Y. Acad. Sci. 1396 (1), 222–235.
Sui, J., Rotshtein, P., Humphreys, G.W., 2013. Coupling social attention to the self forms a network for personal significance. Proceed. Nat. Acad. Sci. 110 (19), 7607–7612.
Tacikowski, P., Nowicka, A., 2010. Allocation of attention to self-name and self-face: an ERP study. Biol. Psychol. 84 (2), 318–324.
Vabalas, A., Gowen, E., Poliakoff, E., Casson, A.J., 2019. Machine learning algorithm validation with a limited sample size. PLoS ONE 14 (11), e0224365.
Wang, C., Oyserman, D., Liu, Q., Li, H., Han, S., 2013. Accessible cultural mind-set modulates default mode activity: evidence for the culturally situated brain. Soc. Neurosci. 8 (3), 203–216.
Wang, Q., Hu, K., Wang, M., Zhao, Y., Liu, Y., Fan, L., Liu, B., 2021a. Predicting brain age during typical and atypical development based on structural and functional neuroimaging. Hum. Brain Mapp. 42 (18), 5943–5955.
Wang, Y., Qin, Y., Li, H., Yao, D., Sun, B., Gong, J., Dai, Y., Wen, C., Zhang, L., Zhang, C., 2021b. Identifying internet addiction and evaluating the efficacy of treatment based on functional connectivity density: a machine learning study. Front. Neurosci. 15, 737.
Wójcik, M.J., Nowicka, M.M., Bola, M., Nowicka, A., 2019. Unconscious detection of one's own image. Psychol. Sci. 30 (4), 471–480.
Wolff, A., Di Giovanni, D.A., Gómez-Pilar, J., Nakao, T., Huang, Z., Longtin, A., Northoff, G., 2019. The temporal signature of self: temporal measures of resting-state EEG predict self-consciousness. Hum. Brain Mapp. 40 (3), 789–803.
Worden, R., Bennett, M.S., Neacsu, V., 2021. The thalamus as a blackboard for perception and planning. Front. Behav. Neurosci. 15, 633872.
Woźniak, M., Knoblich, G., 2019. Self-prioritization of fully unfamiliar stimuli. Q. J. Experiment. Psychol. 72 (8), 2110–2120.
Woźniak, M., Schmidt, T.T., Wu, Y.h., Blankenburg, F., Hohwy, J., 2022. Differences in working memory coding of biological motion attributed to oneself and others. Hum. Brain Mapp. 43 (12), 3721–3734.
Yan, C.G., Wang, X.D., Zuo, X.N., Zang, Y.F., 2016. DPABI: data processing & analysis for (resting-state) brain imaging. Neuroinformatics 14 (3), 339–351.
Yankouskaya, A., Humphreys, G., Stolte, M., Stokes, M., Moradi, Z., Sui, J., 2017. An anterior–posterior axis within the ventromedial prefrontal cortex separates self and reward. Soc. Cogn. Affect. Neurosci. 12 (12), 1859–1868.
Yankouskaya, A., Sui, J., 2022. Self-prioritization is supported by interactions between large-scale brain networks. Eur. J. Neurosci. 55 (5), 1244–1261.
Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 106 (3), 1125–1165.
Yeung, A.W.K., More, S., Wu, J., Eickhoff, S.B., 2022. Reporting details of neuroimaging studies on individual traits prediction: a literature survey. Neuroimage, 119275.
Yin, S., Bi, T., Chen, A., Egner, T., 2021. Ventromedial prefrontal cortex drives the prioritization of self-associated stimuli in working memory. J. Neurosci. 41 (9), 2012–2023.
Ying, X., 2019. An overview of overfitting and its solutions. J. Phys.: Conference Series.
Yip, S.W., Scheinost, D., Potenza, M.N., Carroll, K.M., 2019. Connectome-based prediction of cocaine abstinence. Am. J. Psychiatry 176 (2), 156–164.
Yu, J., Rawtaer, I., Fam, J., Feng, L., Kua, E.-H., Mahendran, R., 2020. The individualized prediction of cognitive test scores in mild cognitive impairment using structural and functional connectivity features. Neuroimage 223, 117310.
Zhang, C., Dougherty, C.C., Baum, S.A., White, T., Michael, A.M., 2018. Functional connectivity predicts gender: evidence for gender differences in resting brain connectivity. Hum. Brain Mapp. 39 (4), 1765–1776.
Zhang, J., Northoff, G., 2022. Beyond noise to function: reframing the global brain activity and its dynamic topography. Commun. Biol. 5 (1), 1350.