



Published in final edited form as:

*Neuroimage*. 2018 February 15; 167: 11–22. doi:10.1016/j.neuroimage.2017.11.010.

## Connectome-based predictive modeling of attention: Comparing different functional connectivity features and prediction methods across datasets

Kwangsun Yoo<sup>1</sup>, Monica D Rosenberg<sup>1</sup>, Wei-Ting Hsu<sup>1</sup>, Sheng Zhang<sup>2</sup>, Chiang-Shan R Li<sup>2,3,4</sup>, Dustin Scheinost<sup>5</sup>, R Todd Constable<sup>4,5,6</sup>, and Marvin M Chun<sup>1,3,4</sup>

<sup>1</sup>Department of Psychology, Yale University, New Haven, CT 06520, USA

<sup>2</sup>Department of Psychiatry, Yale School of Medicine, New Haven, CT 06520, USA

<sup>3</sup>Department of Neuroscience, Yale School of Medicine, New Haven, CT 06520, USA

<sup>4</sup>Interdepartmental Neuroscience Program, Yale University, New Haven, CT 06520, USA

<sup>5</sup>Department of Radiology and Biomedical Imaging, Yale School of Medicine, New Haven, CT 06520, USA

<sup>6</sup>Department of Neurosurgery, Yale School of Medicine, New Haven, CT 06520, USA

### Abstract

Connectome-based predictive modeling (CPM; Finn et al., 2015; Shen et al., 2017) was recently developed to predict individual differences in traits and behaviors, including fluid intelligence (Finn et al., 2015) and sustained attention (Rosenberg et al., 2016a), from functional brain connectivity (FC) measured with fMRI. Here, using the CPM framework, we compared the predictive power of three different measures of FC (Pearson's correlation, accordance, and discordance) and two different prediction algorithms (linear and partial least square [PLS] regression) for attention function. Accordance and discordance are recently proposed FC measures that respectively track in-phase synchronization and out-of-phase anti-correlation (Meskaldji et al., 2016). We defined connectome-based models using task-based or resting-state FC data, and tested the effects of (1) functional connectivity measure and (2) feature-selection/prediction algorithm on individualized attention predictions. Models were internally validated in a training dataset using leave-one-subject-out cross-validation, and externally validated with three independent datasets. The training dataset included fMRI data collected while participants performed a sustained attention task and rested (N=25; Rosenberg et al., 2016a). The validation datasets included: 1) data collected during performance of a stop-signal task and at rest (N=83, including 19 participants who were administered methylphenidate prior to scanning; Rosenberg et al., 2016b; f al., 2014a), 2) data collected during Attention Network Task performance and rest (N=41, Rosenberg et al., in press), and 3) resting-state data and ADHD symptom severity from the ADHD-200 Consortium (N=113; Rosenberg et al., 2016a). Models defined using all combinations of functional

---

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

connectivity measure (Pearson's correlation, accordance, and discordance) and prediction algorithm (linear and PLS regression) predicted attentional abilities, with correlations between predicted and observed measures of attention as high as 0.9 for internal validation, and 0.6 for external validation (all  $p$ 's < 0.05). Models trained on task data outperformed models trained on rest data. Pearson's correlation and accordance features generally showed a small numerical advantage over discordance features, while PLS regression models were usually better than linear regression models. Overall, in addition to correlation features combined with linear models (Rosenberg et al., 2016a), it is useful to consider accordance features and PLS regression for CPM.

## Keywords

functional connectivity; predictive model; attention; partial least square regression

---

## Introduction

Since the concept of functional brain connectivity was introduced, studies in the field have primarily focused on revealing group-wise organizational patterns or differences with brain disorders (Fox and Raichle, 2007), while individual variance was considered to be noisy and unreliable. Recent studies, however, have shown that individual variance is informative, and functional connectivity research is moving towards personalized analyses (Dubois and Adolphs, 2016). It is now evident that functional connectivity patterns are unique for individuals (Finn et al., 2015) and stable across sessions and cognitive states such as task performance and rest (Cole et al., 2014; Finn et al., 2015, 2017; Noble et al., 2017a). Furthermore, individuals' unique patterns of functional brain connectivity contain information about traits, behaviors, and clinical diagnoses (Arbabshirani et al., 2017; Rosenberg et al., 2016a, b). Thus, functional connectivity patterns measured during task performance and rest can provide reliable measures of individual differences in behavior.

Connectome-based predictive modeling (CPM) was recently developed to predict cognitive functions in novel individuals (Finn et al., 2015; Rosenberg et al., 2016a; Shen et al., 2017). Briefly, CPM takes as input whole-brain functional connectivity data, and selects relevant functional connections to predict a behavior or trait that varies across individuals. We demonstrated that this intuitive method can accurately predict individual cognitive and attention function (Finn et al., 2015; Rosenberg et al., 2016a, b). Rosenberg et al. (2016a) showed that a predictive model defined with CPM successfully predicts individuals' sustained attention task performance, a proxy for their overall ability to maintain focus (Rosenberg et al., 2016a). Importantly, this model, the sustained attention connectome-based predictive model or saCPM, predicted other measures of attention in three completely independent data sets. First, the saCPM generalized to predict children's and adolescents' ADHD symptom severity, suggesting that a common functional mechanism underlies performance on lab-based attention tasks and real-world attention function across development (Rosenberg et al., 2016a). In addition, the saCPM generalized to predict performance on a stop-signal task (Rosenberg et al., 2016b) and Attention Network Task (ANT) (Rosenberg et al., in press), both of which require attention for successful performance. The saCPM also characterizes differences in functional connectivity and

behavior that arise due to administration of methylphenidate, a common ADHD treatment that enhances focused attention (Rosenberg et al., 2016b).

In CPM, features are functional connections, or edges, between every pair of regions, or nodes, in the brain. From all possible edges in the brain, a subset is selected for model building based on their relationship with a behavioral measure of interest. In our previous work, we applied a significance threshold (e.g.,  $p < 0.01$ ) to select edges that were positively and negatively related to behavior across subjects. We then calculated a summary statistic, “network strength”, by summing all selected edges. Finally, we constructed a linear regression model with this statistic to predict a behavioral score.

There are several ways to define functional connectivity. The most popular measure is Pearson’s correlation, which quantifies the linear interaction between brain regions. Our group has used this definition for all CPM studies to date (reviewed in Rosenberg et al., 2017; Shen et al., 2017). Pearson’s correlation, however, can be highly affected by small fluctuations that might not be meaningful in fMRI given the low contrast-to-noise ratio. Additionally, Pearson’s  $r$  cannot quantify non-stationary fluctuations (Fox and Greicius, 2010; Hutchison et al., 2013; Meskaldji et al., 2016a; Welvaert et al., 2013). A recent study introduced new measures of functional connectivity – accordance and discordance – which quantify in-phase synchronization and out-of-phase anti-correlation separately, and are calculated based on extreme values in time-series, ignoring low amplitude fluctuations (Meskaldji et al., 2015). Accordance measures the degree to which two signals are co-activated and co-deactivated over the threshold (in-phase synchronization). In contrast, discordance measures the degree to which two signals are oppositely activated or deactivated (out-of-phase anti-correlation), capturing instances when one signal is activated and another is deactivated at the same time. Meskaldji and colleagues proposed that these measures, which are able to capture non-stationary interactions, are more reliable measures of functional connectivity than Pearson’s correlation and more sensitive to detecting differences between groups.

In their work using accordance and discordance to predict individual differences in memory, Meskaldji and colleagues also employed a different method than CPM for feature selection and prediction: partial least square (PLS) regression. PLS regression is a statistical method related to principal component analysis (PCA) and multiple linear regression (Wold et al., 1984), and works by finding a linear relationship between predictor and response variables. This is particularly useful when the number of predictor variables (in this case, edges) is much higher than the number of observations (in this case, participants). PLS regression is based on the same technique used by PCA, representing the predictor variables with a fewer number of latent components. However, in PLS regression, the predictor variables are projected to a new space of components with regard to response variables. Thus, one major strength of PLS regression is that fewer components are required to adequately explain variance in response variables (Mahesh et al., 2015; Wentzell and Vega Montoto, 2003). In Meskaldji et al. (2016a), the PLS regression models using accordance and discordance successfully predicted individual performance on a memory task. The study concluded that discordance, a measure of anti-correlation, contains key information about memory and other cognitive functions (Meskaldji et al., 2016a, b).

An open question is how different connectivity features and feature-selection/prediction algorithms contribute to the success of predictive network models. Investigating the ability of different combinations of features and algorithms to predict behavior will provide insight into the degree to which measures such as accordance and discordance relate to cognitive abilities beyond memory, and inform future work using functional connectivity patterns to make individualized behavioral predictions.

Thus, in the current study, we compare all combinations of three measures of functional connectivity (Pearson's correlation, accordance and discordance) and two feature selection and prediction algorithms (linear and PLS regression) in behavioral predictions, using attentional abilities (measured by ADHD symptom severity and performance on three attention tasks) as a case study. In previous work we demonstrated that linear regression-based CPM is a straightforward but powerful method for accurately predicting an individual's ability to sustain attention. Here, we seek to answer two questions with broad implications for connectome-based predictions: 1) Are accordance and discordance better predictors of attention than Pearson's correlation? 2) Is PLS regression a better predictive algorithm than linear regression for attention? In addressing these questions in the context of attention, we hope to identify broad trends in the performance of different connectome-based predictive modeling approaches that may extend to other abilities and behaviors. In addition, we advocate the importance of considering both internal and external validity. Internal validity refers to the ability of a model to generalize to unseen subjects within a single data set, tested, for example, with leave-one-out cross-validation (LOOCV). Although internal validation is a common standard, it does not ensure that models generalize to new datasets. Testing models' external validity by applying them to completely independent samples ensures that they generalize beyond a single group of participants and minimizes concerns of overfitting (Gabrieli et al., 2015). Further, testing whether models generalize across populations, clinical states, imaging acquisition systems, etc., is important for establishing both theoretical significance and translational utility.

## Methods

### Internal validation: gradual-onset continuous performance task (task-based fMRI and resting-state fMRI)

**Participants**—Thirty-one participants were recruited from Yale University and the surrounding community. Participants performed a sustained attention task, the gradual-onset continuous performance task (gradCPT, Esterman et al., 2013; Rosenberg et al., 2013), during fMRI scanning. Data from six individuals were excluded from analysis due to excessive head motion during fMRI. All participants gave written informed consent in accordance with the Yale University Human Subjects Committee. Detailed information on participants and exclusion criteria can be found in our previous work (Rosenberg et al., 2016a).

**Experimental paradigm**—In the gradCPT, stimuli were grayscale images of city and mountain scenes presented at the center of the screen. On each trial, an image gradually transitioned from one to the next over the course of 800 ms. Participants were instructed to

press a button to city scenes, which randomly occurred 90% of the time, and to withhold response to mountain scenes.

Task runs consisted of four three-minute blocks of the gradCPT interleaved with three 30-second break periods during which a fixation circle appeared in the center of the screen. After each rest break, a dot replaced the circle for two seconds to indicate that the task was about to begin again. Each task run began with eight seconds of fixation, which were excluded from analyses. A total of three task runs (twelve blocks of gradCPT) were administered. During breaks and two rest runs, participants were instructed to attend to the fixation circle in the center of the screen.

Sensitivity ( $d'$ ) was used to measure task performance. For each task block,  $d'$  was calculated as  $z(\text{hit rate}) - z(\text{false alarm rate})$ . An overall  $d'$  value was calculated for each participant by averaging  $d'$  across blocks.

**Image parameters and preprocessing**—MRI data were collected at the Yale Magnetic Resonance Research Center (3T Siemens Trio TIM system equipped with a 32 channel head coil). Resting and task fMRI data were acquired using multiband echo-planar imaging (EPI) with the following parameters: repetition time (TR) = 1,000 ms, echo time (TE) = 30 ms, flip angle = 62°, acquisition matrix = 84 × 84, in-plane resolution = 2.5 mm<sup>2</sup>, 51 axial-oblique slices parallel to the anterior commissure–posterior commissure (AC–PC) line, slice thickness = 2.5 mm, multiband 3, acceleration factor = 2. MPRAGE parameters were as follows: TR = 2530 ms, TE = 2.77 ms, flip angle = 7°, acquisition matrix = 256 × 256, in-plane resolution = 1.0 mm<sup>2</sup>, slice thickness = 1.0 mm, 176 sagittal slices. Two 6-min resting scans were acquired before and after the three task runs. Each task run was 13:44 min long. Volumes collected during rest breaks and the 6 s following them were excluded from analysis. The 2D T1-weighted structural image was acquired with the same slice prescription as the EPI images.

fMRI data were preprocessed using BioImage Suite and custom MATLAB (Mathworks) scripts. Images were motion-corrected using Statistical Parametric Mapping 8 (SPM8, <http://www.fil.ion.ucl.ac.uk/spm/software/spm8/>). Then, several covariates of no interest were regressed from the data, including linear and quadratic drift, cerebrospinal fluid (CSF) signal, white matter (WM) signal, global signal, and a 24-parameter motion model (6 motion parameters, 6 temporal derivatives, and their squares). Data were temporally smoothed with a zero mean unit variance Gaussian filter.

fMRI runs with excessive head motion, defined a priori as > 2 mm translation or > 3° rotation, were excluded from analysis. As a result, one resting run from two participants and one task run from 5 participants were excluded.

### **External validation 1: Methylphenidate and stop-signal task (task-based fMRI and resting-state fMRI)**

**Participants**—We used a dataset of healthy adults originally collected by Farr and colleagues (Farr et al., 2014a, b) and analyzed in our previous work (Rosenberg et al., 2016b). These data included 24 drug-treated participants and 70 unmedicated controls. The

drug-treated subjects were given a single 45 mg dose of methylphenidate about 40 minutes before fMRI scanning and were notified prior to the study that a drug or a placebo would be given. The control subjects did not receive any drug and were not given any drug-related instructions. Detailed participant information can be found in previous studies (Farr et al., 2014a, b). Data exclusion was performed as described in Rosenberg et al., 2016b.

**Experimental paradigm**—Subjects performed a stop-signal task in the scanner (Farr et al., 2014a). Each task trial began with the appearance of a dot in the center of a screen. After a random interval of one to five seconds, the dot changed to a circle, representing a “go” signal. Participants were instructed to press a button as quickly as possible when this go signal appeared. Once the button was pressed, or after one second if no response was detected, the circle disappeared. Trials were separated by 2-second intervals.

About 25% of trials were stop trials. In stop trials, an “X,” the stop signal, replaced the go signal after it appeared. Participants were instructed to withhold response whenever they saw a stop signal. The stop-signal delay (time between the go and stop-signal onsets) was stair-cased within-subject so that each participant was able to withhold response on approximately half of stop trials. As described previously (Rosenberg et al., 2016b), task performance was assessed with go rate, the percent of the time a subject correctly responded to a go signal.

**Image parameters and preprocessing**—MRI data were collected at the Yale Magnetic Resonance Research Center using the same model as that in the internal validation. Images were acquired using an EPI sequence with the following parameters: TR = 2,000 ms, TE = 25 ms, flip angle = 85°, field of view (FOV) = 220 × 220 mm, acquisition matrix = 64 × 64, 32 axial slices parallel to the AC–PC line; slice thickness = 4 mm (no gap). Four 9:50-min task runs and one resting scan of the same length were acquired. A high-resolution T1-weighted structural gradient echo scan was acquired with the following parameters: TR = 2530 ms, TE = 3.66 ms, flip angle = 7°, FOV = 256 × 256 mm, acquisition matrix = 256 × 256, 176 slices parallel to the AC–PC line, slice thickness = 1 mm.

fMRI data were preprocessed as described in the gradCPT section above. fMRI runs with excessive head motion, defined a priori as > 2 mm translation, > 3° rotation, or > 0.15 mm mean frame-to-frame displacement (FD) were excluded from analysis. When an fMRI run had excessive motion near the start or end of the run, volumes were removed as described in Rosenberg et al. 2016b. A total of 88 participants (65 controls) remained after exclusion. Of these, 76 (56 controls) had a rest run and 86 (64 controls) had at least one task run.

## **External validation 2: Attention Network Task (task-based fMRI and resting-state fMRI)**

**Participants**—Fifty-six subjects participated in an fMRI experiment in which they were instructed to rest then perform ANT (Rosenberg et al., in press). Eight subjects were excluded because of excessive head motion in three or more task runs (defined a priori as > 2 mm translation, > 3° rotation, or > 0.15 mm mean FD) and four were excluded for missing data from three or more task runs. Three were excluded from further analysis because they had previously participated in the gradCPT study. The remaining 41 participants were right-handed and had normal or corrected to normal vision (29 females, mean age 23.9 years



[range 18–37]). All participants were paid for participating and gave written informed consent in accordance with the Yale University Human Subjects Committee.

**Experimental paradigm**—Subjects completed the ANT during scanning. Experimental procedures, described in previous work (Rosenberg et al., in press), exactly replicated Fan et al. (2005) except for increased inter-trial intervals (ITIs) (Fan et al., 2005). ANT trials began with a 200-ms cue period. Trials were equally divided between center-cue, spatial-cue, and no-cue trials. On central-cue trials, an asterisk appeared in the center of the screen to alert subjects to an upcoming target. On spatial-cue trials, an asterisk appeared above or below the center fixation cross to indicate with 100% validity where on the screen the upcoming target would appear. On no-cue trials, no cue appeared at all. After a variable period of time (0.3–11.8 seconds; mean = 2.8 seconds), five horizontal arrows appeared above or below the central fixation cross. Participants were instructed to press a button indicating whether the center arrow pointed to the left or to the right.

In half of trials (congruent trials), the central target arrow pointed in the same direction as the flanking distractor arrows. In the other half of trials (incongruent trials), the target pointed in the opposite direction. After a button press (or after two seconds if no response was detected), the arrows disappeared and a variable ITI (5–17 seconds; mean = 8 seconds) began.

Each of the six ANT runs consisted of two buffer trials followed by 36 task trials. Each run included six examples of the six trial types (3 Cue Types: central-cue vs. spatial-cue vs. no-cue  $\times$  2 Target Types: congruent vs. incongruent). Trial order was counterbalanced within and across runs. ANT performance was measured with the variability of correct-trial RTs (Rosenberg et al., in press).

**Image parameters and preprocessing**—MRI data were collected at the Yale Magnetic Resonance Research Center on the same system as the gradCPT dataset. Experimental sessions began with a high-resolution anatomical scan, followed by two 6-min resting scans and six 7:05-min ANT runs. Participants were instructed to fixate on a cross, presented in the center of the screen during rest runs, and to respond to stimuli quickly and accurately during task runs. Rest runs included 360 whole-brain volumes and task runs included 425 volumes. Functional and structural MRI scans were acquired with the same parameters as the gradCPT dataset.

FMRI data were preprocessed as described in internal validation. In the final set of 41 participants, runs with excessive head motion, defined a priori as  $> 2$  mm translation,  $> 3^\circ$  rotation, or  $> 0.15$  mm mean FD (Rosenberg et al., 2016a, b), were excluded from analysis. As a result, one task run was excluded from three participants, and two task runs were excluded from one participant. Due to technical issues or insufficient scan time, two additional participants were missing 205 and 81 volumes from one task run, five participants were missing one task run, one participant was missing two task runs, and one participant was missing three task runs. All participants had two rest runs with acceptable levels of motion.

After calculating connectivity matrices (described in the *Network construction* section below), we excluded any edge (measured with Pearson's correlation) in the task or rest matrices correlated with motion across subjects ( $p < 0.05$ ). This edge exclusion step was performed because head motion can confound connectivity analyses and affect prediction results, and some motion parameters were correlated with ANT performance across subjects. This step resulted in the removal of 9,772 out of 35,778 edges (Rosenberg et al., in press).

### External validation 3: ADHD-200 (resting-state fMRI)

**Participants**—As described previously (Rosenberg et al., 2016a), we used a publically available dataset provided by the ADHD-200 Consortium (Consortium, 2012). This dataset includes resting-state fMRI data from children and adolescents with and without ADHD from eight sites. We used data acquired from Peking University because of the large number of subjects and low head motion. The Research Ethics Review Board of the Peking University Institute of Mental Health approved data collection; all participants agreed to participate and parents or guardians of every participant gave informed consent. Detailed information of inclusion criteria, imaging parameters, and procedures are available online ([http://fcon\\_1000.projects.nitrc.org/indi/adhd200/](http://fcon_1000.projects.nitrc.org/indi/adhd200/)). After exclusion for motion, 113 participants (age 8–16) remained (Rosenberg et al., 2016a).

**Image parameters and preprocessing**—fMRI data were corrected for slice-time and motion using SPM5 (<http://www.fil.ion.ucl.ac.uk/spm/software/spm5/>) then iteratively smoothed until the smoothness of any image had a full width half maximum of approximately 6 mm. All further preprocessing steps were performed using BioImage Suite. Nuisance covariates including linear and quadratic drift, CSF signal, WM signal, global signal, and 6 rigid-body motion parameters, were regressed from the data. Global signal regression is necessary to maximize the prediction, according to prior works (Finn et al., 2015; Rosenberg et al., 2016a; Shen et al; 2017) and additional analysis here. Data were temporally smoothed with a zero mean unit variance Gaussian filter. Due to a lack of full coverage, 32 nodes were excluded from the 268-node functional atlas as described previously (Rosenberg et al., 2016a).

### Comparisons of connectome-based predictive models

Twelve variations of CPM were compared in this study. Models used one of three connectivity definitions (Pearson's correlation, accordance, or discordance), one of two features selection and predictive algorithms (linear or PLS regression), and one of two training datasets (gradCPT-based fMRI or resting-state fMRI). All models were validated using LOOCV in the gradCPT dataset and applied to the three external datasets. Of the twelve models, we previously published the model using Pearson's correlation, gradCPT-based fMRI, and linear regression (Rosenberg et al., 2016a, b). Results from that model were replicated in this study for comparison.

We additionally investigated whether a predictive model can be improved by combining features or data. Predictive models were trained with all 3 connectivity features concatenated or with 2 fMRI datasets (task-based and resting-state) combined. To build and validate a model with combining task-based and resting-state fMRI data, we used three datasets



excluding ADHD data (External 3, which only had resting-state data). We found that models with combined features or data did not outperform the task-based model's performance on task-based data (See Supplementary Table S1, S2).

**Network construction**—Nodes were defined using the Shen 268-node functional brain atlas, which covers the cortex, subcortex and cerebellum (Shen et al., 2013). This atlas was transformed from MNI space to individual space with transformations calculated using intensity-based registration algorithms in BioImage Suite. Average time-series were extracted for each node from preprocessed data. For subjects with more than two usable runs, data were concatenated to obtain one set of time-series. Time-series were normalized to zero mean and unit variance. Then, functional connectivity, defined as the relationship between the time-series of every pair of nodes in the Shen atlas, was computed. Three different functional connectivity measures were estimated and separately used for further prediction analysis: Pearson's correlation, as used in previous studies, and accordance and discordance, which were suggested in recent literature (Meskaldji et al., 2015).

To calculate accordance and discordance, time-series were first thresholded with a predefined quantile of  $q = 0.9$  to retain only significant signals and exclude noise. This  $q$ -value means that only  $\sim (1 - q) * 100\%$  of the highest and of lowest signals in time-series would remain (see details in Meskaldji et al., 2015). We tested with a range of threshold and found similar results across different thresholds (See Supplementary Table S3, S4). For each normalized time-series  $z$ , positively thresholded vector  $z^u$  and negatively thresholded vector  $z^l$  are defined such that

$$\forall \text{time point } t, z_t^u = 0, \text{ if } z_t < \Phi^{-1}(q) \text{ and } z_t^u = 1 \text{ otherwise}$$

$$\forall \text{time point } t, z_t^l = 0, \text{ if } z_t > \Phi^{-1}(1 - q) \text{ and } z_t^l = 1 \text{ otherwise.}$$

where  $\Phi^{-1}$  is the inverse cumulative distribution function of the Gaussian distribution.  $z^u$  and  $z^l$  represent activation and deactivation time, respectively. Accordance  $a_{ij}$  and discordance  $d_{ij}$  between nodes  $i$  and  $j$  were given by,

$$a_{ij} = (z_i^u * z_j^u + z_i^l * z_j^l) / \sigma_i \sigma_j$$

$$d_{ij} = (z_i^u * z_j^l + z_i^l * z_j^u) / \sigma_i \sigma_j$$

where  $\sigma_i = \sqrt{z_i^u * z_i^u + z_i^l * z_i^l}$

**Construction of a predictive model using linear or PLS regression**—We previously demonstrated that CPM (Shen et al., 2017) predicts sustained attentional abilities in novel individuals and datasets (Rosenberg et al., 2016a, b). In these studies, Fisher's  $z$ -

transformed Pearson's correlation coefficients were used to measure functional connectivity. Feature selection was performed by relating the strength of each edge to a measure of attention across subjects with linear regression. Summary statistics (network strengths) were calculated by summing edges related to behavior at a statistical threshold (e.g.,  $p < .01$ ), and the statistics were entered into another linear regression predictive model with the sustained attentional ability.

In this study, we investigated whether alternative connectivity definitions, concordance and discordance, and an alternative feature selection and predictive algorithm, PLS regression, could predict attention as well or better than the previous linear regression CPM.

PLS regression is a statistical regression method related to principal component regression (PCR) and multiple linear regression. It is used to find the relationship between variables to make predictions. Compared to multiple linear regression, PLS regression is particularly useful when the number of predictor variables is much higher than the number of observations. PLS regression utilizes a similar concept as PCR, representing the input variables with fewer components. However, it is considered to be more cost-effective (requiring fewer number of components for prediction) compared to PCR while achieving equal if not better results (Wentzell and Vega Montoto, 2003). The general form of PLS regression is

$$X = TP^T + E$$

$$Y = UQ^T + F$$

where T and U are projections of X and Y respectively, P and Q are loadings, and E and F are the residual errors. In this study, X corresponds to the set of connectivity matrices and Y corresponds to the measures of attention. X and Y are decomposed so as to have the maximized covariance between T and U. We used SIMPLS algorithm to solve PLS regression as implemented in MATLAB R2016b (de Jong, 1993).

### **Validations of the predictive models: internal and external validations**

Predictive models were constructed and tested in the internal dataset using LOOCV and validated in three independent datasets. For internal validation, a predictive model was trained on data from 24 gradCPT subjects and tested with the left-out 25<sup>th</sup> subject. This was done iteratively for all 25 subjects. Predictive models trained in the entire gradCPT dataset were applied to each external validation dataset.

Prediction performance was assessed by correlating predicted and observed behavioral scores. To measure attention, we used  $d'$  (gradCPT/internal validation), go response rate (stop-signal task with methylphenidate/external validation 1), RT variability (ANT/external validation 2), and ADHD symptom severity (ADHD-200/external validation 3). Correlations between predicted and observed scores in external validation 2 and 3 were expected to be negative because of the behavioral/clinical scores used.

### Similarity among measures of functional connectivity

To investigate whether Pearson's correlation, accordance, and discordance provide similar measures of functional brain organization, we compared whole-brain connectivity patterns computed with each measure. That is, for every subject in the gradCPT dataset, we generated six functional connectivity matrices: A task and rest matrix computed using Pearson's correlation, accordance, and discordance. Similarity was then quantified as the Pearson's spatial correlation between each pair of task matrices, each pair of rest matrices, and each pair of matrices computed from the same measure (e.g., task-based and resting-state accordance matrices).

### Anatomical distribution of edges with significant PLS regression coefficient

Instead of only including certain features in prediction, PLS regression includes all features but assigns each a different weight. We performed  $10^6$  permutations to find the edges with statistically significant weights in PLS regression. In each permutation, behavioral scores ( $d'$ ) were randomly shuffled across subjects and PLS regression was performed. PLS regression beta coefficients for every edge were obtained from each permutation. For each edge, significance was determined by whether its real beta value differed (two-tailed  $p < 0.05$ ) from the empirical distribution acquired from  $10^6$  permutations. Significance testing was performed separately for each measure of functional connectivity (Pearson's correlation, accordance, and discordance). This analysis was performed using task-based data from the gradCPT dataset.

### Comparison between edges in PLS regression and edges in linear regression models

We investigated the relationship between important edges in the PLS predictive model and those in the saCPM's high-attention network (HAN) and low-attention network (LAN), defined in previous work (Rosenberg et al., 2016a). As a first approach, we compared PLS beta weights of edges that appeared in the HAN, LAN, or neither. As a second approach, we calculated the overlap between the important edges in a PLS predictive model and those in the HAN or LAN.

## Results

Each model's prediction performance was assessed by correlating predicted and observed behavioral scores (e.g.,  $d'$  for gradCPT) across individuals.

### Internal validation

**Predictions from task-based fMRI**—Twelve models (3 functional connectivity measures [Pearson's correlation vs. accordance vs. discordance]  $\times$  2 prediction algorithms [linear vs. PLS regression]  $\times$  2 types of data [task-based or resting-state]) were trained using 24 gradCPT subjects' task data and tested on the left-out 25<sup>th</sup> subject. Of models tested on the left-out subject's task data, PLS regression models trained on gradCPT-based fMRI provided numerically higher predictions. Predictions were significant for PLS regression models using all three connectivity measures (Pearson's  $r = 0.884$ ,  $p = 4.86 \times 10^{-9}$ ; accordance  $r = 0.908$ ,  $p = 3.54 \times 10^{-10}$ ; discordance:  $r = 0.893$ ,  $p = 1.93 \times 10^{-9}$ ). All three correlations were significant at  $p < 10^{-6}$  and larger than the largest value from the null

distribution obtained from  $10^6$  permutations (Figure 1 top and Table 1). The predictive power of all twelve models can be found in Figure 1 and Table 1.

**Predictions from resting-state fMRI**—Models trained on gradCPT task data significantly predicted attention from functional connectivity observed at rest (Pearson's  $r = 0.527$ , parametric  $p = 6.78 \times 10^{-3}$  [empirical  $p = 1.36 \times 10^{-3}$ ]; accordance:  $r = 0.509$ ,  $p = 9.29 \times 10^{-3}$  [ $p = 1.46 \times 10^{-3}$ ]; discordance  $r = 0.554$ ,  $p = 4.05 \times 10^{-3}$  [ $p = 5.67 \times 10^{-4}$ ]; empirical  $p$  values were based on  $10^6$  permutations). The performance of all twelve models can be found in Figure 1 and Table 1.

### External validations

Prediction performances of all 12 models for all three external validations are also represented in Figure 1 and Table 1. To summarize the performance of all models, six PLS regression predictive models achieved accurate predictions in all external validations ( $p < 0.05$ , Table 1). On average, PLS models trained on resting-state fMRI provided the best prediction when testing on external resting datasets (Figure 2D). Similarly, PLS models trained on task-based fMRI provided numerically higher prediction accuracy than other models when tested on external task-based fMRI.

### Similarity among different functional connectivity feature calculations

We assessed similarities among the three ways of calculating functional connectivity features: Pearson's correlation, accordance, and discordance. The patterns of the three measures were very similar (mean similarity  $> 0.7$  from task fMRI and  $> 0.6$  from resting fMRI, Figure 3). Pearson's correlation and accordance were the most similar ( $\sim 0.9$ , 1<sup>st</sup> & 4<sup>th</sup> pairs in Figure 3), and accordance and discordance were the least similar ( $\sim 0.7$ , 3<sup>rd</sup> & 6<sup>th</sup> pairs in Figure 3). Each measure was also shown to be consistent (similarity over 0.5) between task and resting-state (7~9<sup>th</sup> pairs in Figure 3).

### Anatomical distribution of edges with significant PLS regression coefficients

The following results are based on the three PLS predictive models trained on gradCPT-based fMRI and using Pearson's correlation, accordance, and discordance.

We investigated the anatomical distribution of edges with significant PLS regression coefficient weights (two-tailed  $p < 0.05$  based on  $10^6$  permutations; Figure 4A & Figure 5). Interestingly, many edges with negative weights (i.e., edges whose strength contributes to predictions of worse attention) were found to be inter-hemispheric, especially bilateral temporal and bilateral cerebellar edges (Figure 4A). In the left hemisphere, intra-temporal and intra-cerebellar connectivity also consistently had significant weights (Figure 4B). In contrast, edges with significant positive weights were generally distributed across the brain (Figure 4A), and connectivity within the right occipital lobe and between right motor and left occipital cortex consistently predicted better attention (Figure 4B).

We then directly investigated the number of edges that had significant PLS regression coefficients in models using different connectivity definitions (Figure 4B, C). We found eight common edges with significant positive weights and six common edges with

significant negative weights. Six positive-weight edges constituted a connected set through left occipital area, right cerebellum and right motor area. In contrast, the six negative-weight edges were distributed and included connections within bilateral temporal areas, within left the cerebellum and between left prefrontal and right parietal cortex. Two or more edges within the same pair of lobes were merged into one line (Figure 4B).

**Comparison between edges in the PLS regression model and edges in the linear regression model**—We next investigated the relationship between the saCPM (identified using linear regression in Rosenberg et al., 2016a) and the networks identified using PLS regression (Figure 6). First, we compared the PLS weights of edges present in the saCPM’s high-attention network (HAN; 757 edges whose strength predicts better attention) and low-attention network (LAN; 630 edges whose strength predicts worse attention) with the weights of edges not included in either network (Figure 6A). Next, we counted the number of edges with significant PLS weights that also appeared in the HAN or LAN (Figure 6B). We found that the PLS regression model using Pearson’s correlation included 51 edges with significant positive weights and 28 edges with significant negative weights. All 51 positive edges were included in the HAN, and all 28 edges were included in the LAN. For PLS models based on accordancy, 46 of 47 edges with significant positive weights were included in the HAN, and 75 of 79 edges with significant negative weights were included in the LAN. For PLS models based on discordance, 57 of 59 edges with significant positive weights were included in the HAN, and 21 of 25 edges with significant negative weights were included in the LAN.

## Discussion

Every person has a unique pattern of functional brain connectivity that contains information about their cognitive abilities (Finn et al., 2015; Rosenberg et al., 2016a). A recent model based on these unique connectivity patterns, the sustained attention connectome-based predictive model (saCPM), used linear regression and functional connectivity matrices created using Pearson’s correlation and yielded strong predictions of people’s ability to attend in multiple contexts, and generalized across multiple independent datasets (reviewed in Rosenberg et al., 2017).

The current study tested ways to improve CPM performance across multiple data sets and cognitive states. Accordingly, we built models using two feature-selection/prediction algorithms (linear and PLS regression) using three functional connectivity measures (Pearson’s correlation, accordancy, and discordance) observed during two cognitive states (task engagement and quiet rest) to predict attention. We assessed model performance using LOOCV as well as cross-dataset validation. In general, results suggest that CPM can be improved with more sensitive connectivity measures and more complicated predictive algorithms.

Comparisons of models built with linear or PLS regression indicate that for our attention data, PLS regression-based models offered a slight numerical advantage. Especially, PLS regression models trained on rest data outperformed linear regression models when applied to rest data. Although linear and PLS regression models differ in several respects, feature

selection is one salient distinction. While PLS regression models assign every functional connection a weight for prediction, linear regression models select a subset of functional connections for prediction. If some selected features are not reliable (for example, across sessions or subjects), this method might not be robust. To predict behavior in independent datasets, which differ along many dimensions that can affect functional connectivity, such as demographics, cognitive state, clinical state, and MRI-related factors (Choe et al., 2015; Chou et al., 2012; Zuo and Xing, 2014), it is important to minimize unreliable features. Thus, PLS regression models may be advantageous for cross-dataset prediction. Not tested here, canonical correlation analysis is another method to investigate relationships between two sets of variables (Smith et al., 2015).

Accordance and discordance are recently developed measures of functional connectivity, and have been applied to predict long-term memory function in patients with mild cognitive impairment (Meskaldji et al., 2016a). Meskaldji and colleagues demonstrated an advantage for discordance features in predicting MCI, but in the current study the discordance-based predictive model did not outperform the other two connectivity measures in predicting attention. Instead, accordance and Pearson's correlation generally provided more accurate predictions. For linear regression models defined with task fMRI, accordance was most predictive of attention in four out of the seven tested datasets (gradCPT task and rest, stop-signal rest, and ADHD-200 rest). This may indicate that synchronized interaction among brain regions plays a key role in attention (Buschman and Miller, 2007; Fair et al., 2010; Gregoriou et al., 2009; Just et al., 2004; Killory et al., 2011; Konrad and Eickhoff, 2010; Rosenberg et al., 2017). However, because different behavioral measures may be more closely related to different functional connectivity features, we recommend that future CPM studies consider Pearson's correlation, accordance, and discordance.

In comparing the similarity of functional connectivity patterns measured with Pearson's correlation, accordance, and discordance, we found that correlation- and accordance-based connectivity patterns were highly similar (mean  $r > 0.9$  within individuals in task-states). Further, the majority of edges with significant PLS weights in models trained on Pearson's correlation matrices also had significant weights in models trained on accordance matrices. This is consistent with reports that Pearson's correlation is primarily driven by a few critical points in time-series (Karahano lu and Van De Ville, 2015; Liu and Duyn, 2013; Scheinost et al., 2016; Tagliazucchi et al., 2012). These critical points are generally the points used in calculating accordance. Hence, both Pearson's correlation and accordance rely on similar time points, explaining the high similarity between correlation-based and accordance-based connectivity patterns. In sum, Pearson's correlation and accordance may reflect common characteristics of the relationship between activity in spatially distinct brain regions, whereas discordance may reflect relatively a unique aspect of these relationships.

The anatomy of predictive networks parallels previously published results (Rosenberg et al., 2016a, b; in press) and underscores the importance of connections distributed across cortical, subcortical, and cerebellar regions in predicting individual differences in attention. To highlight several broad trends, functional connections between occipital, motor, and cerebellar regions were consistently related to stronger attentional abilities, perhaps reflecting efficiency in a circuit necessary for executing motor responses during visual



attention tasks. Intra-temporal, intra-cerebellar, and temporal-cerebellar functional connections, on the other hand, were consistently implicated in poor attention function. In conjunction with frequent observations of altered cerebellar structure and function in ADHD (Castellanos et al., 2002; Castellanos and Proal, 2012; Krain and Castellanos, 2006; Mostofsky et al., 1998), these results provide evidence for an important role of the cerebellum in attention.

Examining the contribution of resting-state networks traditionally implicated in attention - the frontoparietal and default mode networks (Corbetta and Shulman, 2002; Mason et al., 2007; Spreng et al., 2010) - further illuminates the distributed nature of attention process across the brain. Previous work used a “computational lesioning” approach to investigate the importance the frontoparietal and default mode network edges to predict individual differences in attention using the CPM approach. Results showed that removing edges with at least one node in either network had small but non-significant effects of predictive power (Rosenberg et al., 2016a). Furthermore, the frontoparietal and default mode networks were not over-represented in the predictive networks defined here. Although 23.81% of all possible edges in the Shen brain atlas (Finn et al., 2015) involve frontoparietal network nodes, only 4 of the 51 Pearson’s correlation-based edges with a significant positive PLS regression coefficient (7.84%), and only 3 of the 28 edges with a significant negative coefficient (10.71%), involved frontoparietal nodes (leftmost column of Figure 4A). Edges involving default mode network nodes make up 14.39% of all possible Shen atlas edges, but only 3.92% of the edges with significant positive and 10.71% of the edges with significant negative PLS regression components. No within-frontoparietal or within-default mode network connections had significant PLS regression coefficients. Incomplete overlap between the current network models and classic network-based models of attention may arise in part from differences in methodology, namely, our relatively new functional connectivity approach based on individual differences, compared to the conventional functional connectivity and task-activation approach based on group averages. Our work does not conflict with prior work, but rather shows that frontoparietal and default mode network activity contribute to attentional performance in concert with a more broadly distributed network across the whole brain. Different insights from each analysis method should inform future research on the distributed nature of functional brain organizational patterns underlying complex cognitive processes such as attention (Rosenberg et al., 2017).

Regardless of connectivity measure or prediction algorithm, models tested on task-based connectivity patterns performed about twice as well as models tested on resting-state connectivity patterns. One reason that models may not generalize from task to rest as well as they do from one task state to another (e.g., from gradCPT to ANT) is that attention reorganizes brain network topology and functional connection strength (Buschman and Miller, 2007; Cole et al., 2013; Gießing et al., 2013; Mainero et al., 2004; Meehan et al., 2017; Saproo and Serences, 2014). This is evident in our own work, as the correlation between a subject’s functional connectivity patterns observed during gradCPT performance and rest was, at maximum, 0.7 (Figure 3). Another recent study using data from the Human Connectome Project found the within-subject similarity of functional connectivity patterns between different brain states was generally lower than 0.7 (Finn et al., 2017).

A second reason that models may not generalize as well to rest is that resting-state data may be less reliable than task data due to the unconstrained nature of rest runs or simply the fact that researchers tend to collect less resting-state than task data. For example, our gradCPT, ANT, and stop-signal datasets included more than three times as many task as rest volumes. It is evident that a reliability of functional connectivity is dependent on data-length (Finn et al., 2015; Laumann et al., 2015; Noble et al., 2017b; Pannunzi et al., 2017). Recent work investigating the reliability of resting-state functional connectivity showed similarity of 0.55 between connectivity patterns calculated from two ~4 minute long runs obtained on different days, but higher similarity can be obtained with longer resting-state runs (Finn et al., 2017). To eliminate the effect of data-length on the predictive power in the current study, we performed the additional analysis with matching task and rest data-length in two datasets (gradCPT: Internal1 and ANT: External 2) that were collected with the same MR acquisition parameters, including 1 sec TR. We built 12 CPMs using the first 350 volumes of task-based or resting-state fMRI from gradCPT dataset, and tested the models in gradCPT and ANT datasets (See Supplementary Table S5). Overall performance was, as one may expect, lower than the original models which were trained with more data (See Supplementary Table S5). However, task-based models continue to outperform the rest models. Another recent study on multisite reliability reported that the similarity between imaging sessions/sites was lower than 0.75 (Noble et al., 2017a). Other works have reported reproducibility lower than 80% (Havel et al., 2006; Meindl et al., 2009; van de Ven et al., 2004). Thus, most studies have concluded that resting fMRI is significantly reliable, but with considerable variability. Finally, it is important to note that the predictive power of all connectome-based models is constrained by the reliability of the behavioral measure of interest and of fMRI itself, and hence prediction accuracy is limited both theoretically and practically.

Although the CPM variants tested here were constructed in a highly data-driven manner, models did have two adjustable parameters: one for generating connectivity features and the other for constructing PLS regression models. The first parameter, necessary for calculating accordancy and discordancy, is the threshold for binarizing time-series by ignoring low amplitudes considered as noise in fluctuations. Meskaldji and colleagues (2016a) selected this threshold ( $q = 0.8$ ) based on previous work that reported that correlation results were stable under a wide range of a parameter values (Tagliazucchi et al., 2012). In this study, we used a stricter threshold,  $q = 0.9$ , meaning that we calculated accordancy and discordancy from approximately 5% of the highest and of lowest signals in the time-series.

The second free parameter was the number of latent components in our PLS regression model (Li et al., 2002). Because PLS regression works through dimensional reduction, prediction results depend on the number of components used. Components in PLS regression are ordered according to the decreasing amount of variance explained. In this study, we used the first component, which explains the most of variance. Although including additional components should explain more variance, according to additional analyses not included here, it did not produce higher prediction accuracy or better generalizability, while increasing the risk of overfitting to the training sample or fitting to noise. Thus, future work may explore the optimal selection of components for building generalizable connectome-based PLS regression models of behavior.

In conclusion, we compared how well models based on different measures of functional connectivity and different feature-selection/prediction algorithms predicted attention. Based on our findings, we recommend that future connectome-based models be defined over task data when possible, since individual differences in behaviorally relevant circuits and networks are likely magnified by ongoing task performance. These models can then be applied to resting-state data, which is all that is available in many large scale studies. For predicting attention, our meta-analysis did not reveal significant advantages for accordance or discordance features over Pearson correlation features, but it would still be useful to compare these different features in a broader range of cognitive tasks. For feature selection, although the differences were small, PLS regression could be a better choice over linear regression for models applied to resting-state data. Finally, we reiterate the value of validating all predictive models across independent data sets.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This project was supported by the National Institute of Health (NIH) grant MH108591, R01AA021449, and R01DA023248, and by the National Science Foundation grant BCS 1558497 and BCS1309260. S.Z was supported by the NIH grant K25DA040032.

## References

- Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *Neuroimage*. 2017; 145:137–165. DOI: 10.1016/j.neuroimage.2016.02.079 [PubMed: 27012503]
- Buschman TJ, Miller EK. Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. *Science* (80-.). 2007; 315
- Castellanos F, PP L, Sharp W, et al. Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. *JAMA*. 2002; 288:1740–1748. [PubMed: 12365958]
- Castellanos FX, Proal E. Large-scale brain systems in ADHD: Beyond the prefrontal-striatal model. *Trends Cogn. Sci.* 2012; doi: 10.1016/j.tics.2011.11.007
- Choe AS, Jones CK, Joel SE, Muschelli J, Belegu V, Caffo BS, Lindquist MA, van Zijl PCM, Pekar JJ. Reproducibility and Temporal Structure in Weekly Resting-State fMRI over a Period of 3.5 Years. *PLoS One*. 2015; 10:e0140134.doi: 10.1371/journal.pone.0140134 [PubMed: 26517540]
- Chou, Y-h, Panych, LP., Dickey, CC., Petrella, JR., Chen, N-k. Investigation of Long-Term Reproducibility of Intrinsic Connectivity Network Mapping: A Resting-State fMRI Study. *Am. J. Neuroradiol.* 2012; 33
- Cole MW, Bassett DS, Power JD, Braver TS, Petersen SE. Intrinsic and Task-Evoked Network Architectures of the Human Brain. *Neuron*. 2014; 83:238–251. DOI: 10.1016/j.neuron.2014.05.014 [PubMed: 24991964]
- Cole MW, Reynolds JR, Power JD, Repovs G, Anticevic A, Braver TS. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat. Neurosci.* 2013; 16:1348–1355. DOI: 10.1038/nn.3470 [PubMed: 23892552]
- Consortium, T.A.-200. The ADHD-200 Consortium: A Model to Advance the Translational Potential of Neuroimaging in Clinical Neuroscience. *Front. Syst. Neurosci.* 2012; 6:62.doi: 10.3389/fnsys.2012.00062 [PubMed: 22973200]
- Corbetta M, Shulman GL. CONTROL OF GOAL-DIRECTED AND STIMULUS-DRIVEN ATTENTION IN THE BRAIN. *Nat. Rev. Neurosci.* 2002; 3:215–229. DOI: 10.1038/nrn755

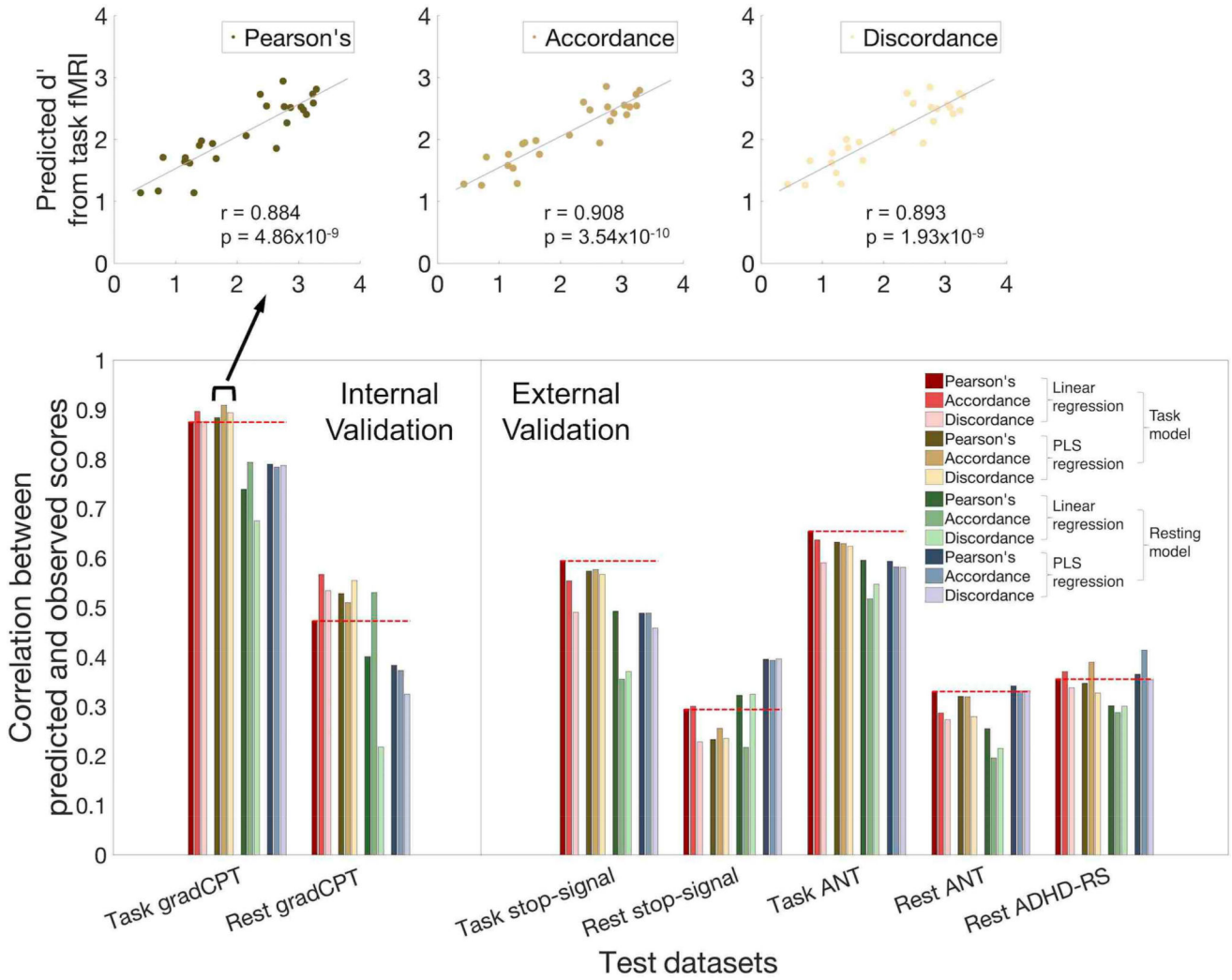
- de Jong S. SIMPLS: An alternative approach to partial least squares regression. *Chemom. Intell. Lab. Syst.* 1993; 18:251–263. DOI: 10.1016/0169-7439(93)85002-X
- Dubois J, Adolphs R. Building a Science of Individual Differences from fMRI. *Trends Cogn. Sci.* 2016; 20:425–443. DOI: 10.1016/j.tics.2016.03.014 [PubMed: 27138646]
- Esterman M, Noonan SK, Rosenberg M, Degutis J. In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cereb. Cortex.* 2013; 23:2712–2723. [PubMed: 22941724]
- Fair DA, Posner J, Nagel BJ, Bathula D, Dias TGC, Mills KL, Blythe MS, Giwa A, Schmitt CF, Nigg JT. Atypical Default Network Connectivity in Youth with Attention-Deficit/Hyperactivity Disorder. *Biol. Psychiatry.* 2010; 68:1084–1091. DOI: 10.1016/j.biopsych.2010.07.003 [PubMed: 20728873]
- Fan J, McCandliss BD, Fossella J, Flombaum JI, Posner MI. The activation of attentional networks. *Neuroimage.* 2005; 26:471–479. DOI: 10.1016/j.neuroimage.2005.02.004 [PubMed: 15907304]
- Farr OM, Hu S, Matuskey D, Zhang S, Abdelghany O, Li CR. The effects of methylphenidate on cerebral activations to salient stimuli in healthy adults. *Exp. Clin. Psychopharmacol.* 2014a; 22:154–165. DOI: 10.1037/a0034465 [PubMed: 24188171]
- Farr OM, Zhang S, Hu S, Matuskey D, Abdelghany O, Malison RT, Li C-SR. The effects of methylphenidate on resting-state striatal, thalamic and global functional connectivity in healthy adults. *Int. J. Neuropsychopharmacol.* 2014b; 17:1177–1191. DOI: 10.1017/S1461145714000674 [PubMed: 24825078]
- Finn ES, Scheinost D, Finn DM, Shen X, Papademetris X, Constable RT. Can brain state be manipulated to emphasize individual differences in functional connectivity? *Neuroimage.* 2017; doi: 10.1016/j.neuroimage.2017.03.064
- Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* 2015; 18:1664–1671. DOI: 10.1038/nn.4135 [PubMed: 26457551]
- Fox MD, Greicius M. Clinical applications of resting state functional connectivity. *Front. Syst. Neurosci.* 2010; 4:19. doi: 10.3389/fnsys.2010.00019 [PubMed: 20592951]
- Fox MD, Raichle ME. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 2007; 8:700–711. DOI: 10.1038/nrn2201 [PubMed: 17704812]
- Gabrieli JDE, Ghosh SS, Whitfield-Gabrieli S. Prediction as a Humanitarian and Pragmatic Contribution from Human Cognitive Neuroscience. *Neuron.* 2015; 85:11–26. DOI: 10.1016/j.neuron.2014.10.047 [PubMed: 25569345]
- Gießing C, Thiel CM, Alexander-Bloch AF, Patel AX, Bullmore ET. Human Brain Functional Network Changes Associated with Enhanced and Impaired Attentional Task Performance. *J. Neurosci.* 2013; 33
- Gregoriou GG, Gotts SJ, Zhou H, Desimone R. High-Frequency, Long-Range Coupling Between Prefrontal and Visual Cortex During Attention. *Science (80-).* 2009; 324
- Havel P, Braun B, Rau S, Tonn J-Ch, Fesl G, Brückmann H, Ilmberger J. Reproducibility of activation in four motor paradigms. *J. Neurol.* 2006; 253:471–476. DOI: 10.1007/s00415-005-0028-4 [PubMed: 16283098]
- Hutchison RM, Womelsdorf T, Allen EA, Bandettini PA, Calhoun VD, Corbetta M, Della Penna S, Duyn JH, Glover GH, Gonzalez-Castillo J, Handwerker DA, Keilholz S, Kiviniemi V, Leopold DA, de Pasquale F, Sporns O, Walter M, Chang C. Dynamic functional connectivity: Promise, issues, and interpretations. *Neuroimage.* 2013; 80:360–378. DOI: 10.1016/j.neuroimage.2013.05.079 [PubMed: 23707587]
- Just MA, Cherkassky VL, Keller TA, Minshew NJ. Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity. *Brain.* 2004; 127:1811–1821. DOI: 10.1093/brain/awh199 [PubMed: 15215213]
- Karahano lu FI, Van De Ville D. Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks. *Nat. Commun.* 2015; 6:7751. doi: 10.1038/ncomms8751 [PubMed: 26178017]

- Killory BD, Bai X, Negishi M, Vega C, Spann MN, Vestal M, Guo J, Berman R, Danielson N, Trejo J, Shisler D, Novotny EJ, Constable RT, Blumenfeld H. Impaired attention and network connectivity in childhood absence epilepsy. *Neuroimage*. 2011; 56:2209–2217. DOI: 10.1016/j.neuroimage.2011.03.036 [PubMed: 21421063]
- Konrad K, Eickhoff SB. Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder. *Hum. Brain Mapp*. 2010; 31:904–916. DOI: 10.1002/hbm.21058 [PubMed: 20496381]
- Krain AL, Castellanos FX. Brain development and ADHD. *Clin. Psychol. Rev*. 2006; 26:433–444. DOI: 10.1016/j.cpr.2006.01.005 [PubMed: 16480802]
- Laumann TO, Gordon EM, Adeyemo B, Snyder AZ, Joo SJ, Chen MY, Gilmore AW, McDermott KB, Nelson SM, Dosenbach NU, Schlaggar BL, Mumford JA, Poldrack RA, Petersen SE. Functional System and Areal Organization of a Highly Sampled Individual Human Brain, 2015. *Neuron*. 2015; 87:657–670. DOI: 10.1016/J.NEURON.2015.06.037 [PubMed: 26212711]
- Li B, Morris J, Martin EB. Model selection for partial least squares regression. *Chemom. Intell. Lab. Syst.* 2002; 64:79–89. DOI: 10.1016/S0169-7439(02)00051-5
- Liu X, Duyn JH. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proc. Natl. Acad. Sci. U. S. A.* 2013; 110:4392–7. DOI: 10.1073/pnas.1216856110 [PubMed: 23440216]
- Mahesh S, Jayas DS, Paliwal J, White NDG. Comparison of Partial Least Squares Regression (PLSR) and Principal Components Regression (PCR) Methods for Protein and Hardness Predictions using the Near-Infrared (NIR) Hyperspectral Images of Bulk Samples of Canadian Wheat. *Food Bioprocess Technol*. 2015; 8:31–40. DOI: 10.1007/s11947-014-1381-z
- Mainero C, Caramia F, Pozzilli C, Pisani A, Pestalozza I, Borriello G, Bozzao L, Pantano P. fMRI evidence of brain reorganization during attention and memory tasks in multiple sclerosis. *Neuroimage*. 2004; 21:858–867. DOI: 10.1016/j.neuroimage.2003.10.004 [PubMed: 15006652]
- Mason MF, Norton MI, Van Horn JD, Wegner DM, Grafton ST, Macrae CN. Wandering minds: the default network and stimulus-independent thought. *Science*. 2007; 315:393–5. DOI: 10.1126/science.1131295 [PubMed: 17234951]
- Meehan TP, Bressler SL, Tang W, Astafiev SV, Sylvester CM, Shulman GL, Corbetta M. Top-down cortical interactions in visuospatial attention. *Brain Struct. Funct*. 2017; :1–19. DOI: 10.1007/s00429-017-1390-6
- Meindl T, Teipel S, Elmouden R, Mueller S, Koch W, Dietrich O, Coates U, Reiser M, Glaser C. Test-retest reproducibility of the default-mode network in healthy individuals. *Hum. Brain Mapp*. 2009; 31:NA–NA. DOI: 10.1002/hbm.20860
- Meskaldji, D-E., Morgenthaler, S., Van De Ville, D. 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI). IEEE; 2015. New measures of brain functional connectivity by temporal analysis of extreme events; p. 26-29.
- Meskaldji D-E, Preti MG, Bolton TA, Montandon M-L, Rodriguez C, Morgenthaler S, Giannakopoulos P, Haller S, Van De Ville D. Prediction of long-term memory scores in MCI based on resting-state fMRI. *NeuroImage Clin*. 2016; 12:785–795. DOI: 10.1016/j.nicl.2016.10.004 [PubMed: 27812505]
- Meskaldji, D-E., Preti, MG., Bolton, T., Montandon, M-L., Rodriguez, C., Morgenthaler, AS., Giannakopoulos, P., Haller, S., Van De Ville, D. 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). IEEE; 2016. Predicting individual scores from resting state fMRI using partial least squares regression; p. 1311-1314.
- Mostofsky SH, Reiss AL, Lockhart P, Denckla MB. Evaluation of Cerebellar Size in Attention-Deficit Hyperactivity Disorder. *J. Child Neurol*. 1998; 13:434–439. DOI: 10.1177/088307389801300904 [PubMed: 9733289]
- Noble S, Scheinost D, Finn ES, Shen X, Papademetris X, McEwen SC, Bearden CE, Addington J, Goodyear B, Cadenhead KS, Mirzakhani H, Cornblatt BA, Olvet DM, Mathalon DH, McGlashan TH, Perkins DO, Belger A, Seidman LJ, Thermenos H, Tsuang MT, van Erp TGM, Walker EF, Hamann S, Woods SW, Cannon TD, Constable RT. Multisite reliability of MR-based functional connectivity. *Neuroimage*. 2017a; 146:959–970. DOI: 10.1016/j.neuroimage.2016.10.020 [PubMed: 27746386]



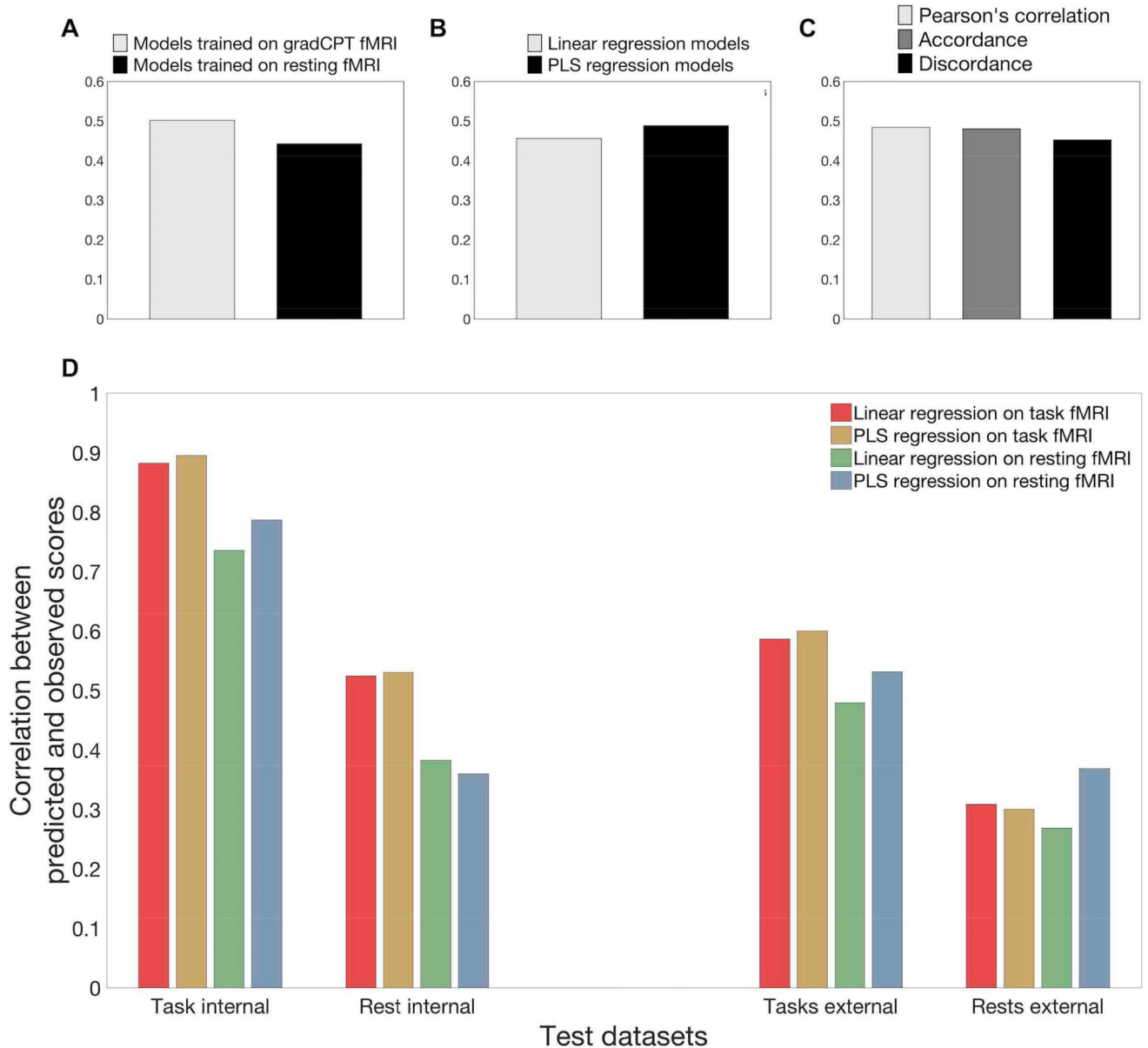
- Noble S, Spann MN, Tokoglu F, Shen X, Constable RT, Scheinost D. Influences on the Test–Retest Reliability of Functional Connectivity MRI and its Relationship with Behavioral Utility. *Cereb. Cortex*. 2017b; :1–15. DOI: 10.1093/cercor/bhx230
- Pannunzi M, Hindriks R, Bettinardi RG, Wenger E, Lisofsky N, Martensson J, Butler O, Filevich E, Becker M, Lochstet M, Kühn S, Deco G. Resting-state fMRI correlations: From link-wise unreliability to whole brain stability. *Neuroimage*. 2017; 157:250–262. DOI: 10.1016/j.neuroimage.2017.06.006 [PubMed: 28599964]
- Rosenberg MD, Finn ES, Scheinost D, Constable RT, Chun MM. Characterizing Attention with Predictive Network Models. *Trends Cogn. Sci*. 2017; 21:290–302. DOI: 10.1016/j.tics.2017.01.011 [PubMed: 28238605]
- Rosenberg MD, Finn ES, Scheinost D, Papademetris X, Shen X, Constable RT, Chun MM. A neuromarker of sustained attention from whole-brain functional connectivity. *Nat. Neurosci*. 2016a; 19:165–171. DOI: 10.1038/nn.4179 [PubMed: 26595653]
- Rosenberg MD, Noonan S, DeGutis J, Esterman M. Sustaining visual attention in the face of distraction: a novel gradual-onset continuous performance task. *Atten. Percept. Psychophys*. 2013; 75:426–439. [PubMed: 23299180]
- Rosenberg MD, Zhang S, Hsu W-T, Scheinost D, Finn ES, Shen X, Constable RT, Li C-SR, Chun MM. Methylphenidate Modulates Functional Network Connectivity to Enhance Attention. *J. Neurosci*. 2016b; 36
- Saproo S, Serences JT. Attention Improves Transfer of Motion Information between V1 and MT. *J. Neurosci*. 2014; 34
- Scheinost D, Tokoglu F, Shen X, Finn ES, Noble S, Papademetris X, Constable RT. Fluctuations in Global Brain Activity Are Associated With Changes in Whole-Brain Connectivity of Functional Networks. *IEEE Trans. Biomed. Eng*. 2016; 63:2540–2549. DOI: 10.1109/TBME.2016.2600248 [PubMed: 27541328]
- Shen X, Finn ES, Scheinost D, Rosenberg MD, Chun MM, Papademetris X, Constable RT. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat. Protoc*. 2017; 12:506–518. DOI: 10.1038/nprot.2016.178 [PubMed: 28182017]
- Shen X, Tokoglu F, Papademetris X, Constable RT. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage*. 2013; 82:403–415. DOI: 10.1016/j.neuroimage.2013.05.081 [PubMed: 23747961]
- Spreng RN, Stevens WD, Chamberlain JP, Gilmore AW, Schacter DL. Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *Neuroimage*. 2010; 53:303–317. DOI: 10.1016/j.neuroimage.2010.06.016 [PubMed: 20600998]
- Tagliazucchi E, Balenzuela P, Fraiman D, Chialvo DR. Criticality in Large-Scale Brain fMRI Dynamics Unveiled by a Novel Point Process Analysis. *Front. Physiol*. 2012; 3:15.doi: 10.3389/fphys.2012.00015 [PubMed: 22347863]
- van de Ven VG, Formisano E, Prvulovic D, Roeder CH, Linden DEJ. Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. *Hum. Brain Mapp*. 2004; 22:165–178. DOI: 10.1002/hbm.20022 [PubMed: 15195284]
- Welvaert M, Rosseel Y, Dyck M, Mathiak K, Mathiak K. On the Definition of Signal-To-Noise Ratio and Contrast-To-Noise Ratio for fMRI Data. *PLoS One*. 2013; 8:e77089.doi: 10.1371/journal.pone.0077089 [PubMed: 24223118]
- Wentzell PD, Vega Montoto L. Comparison of principal components regression and partial least squares regression through generic simulations of complex mixtures. *Chemom. Intell. Lab. Syst*. 2003; 65:257–279. DOI: 10.1016/S0169-7439(02)00138-7
- Wold S, Ruhe IA, Wold H, Dunn WJ Iii. THE COLLINEARITY PROBLEM IN LINEAR REGRESSION. THE PARTIAL LEAST SQUARES (PLS) APPROACH TO GENERALIZED INVERSES\*. *SIAM J. Sci. STAT. Comput*. 1984; 5
- Zuo X-N, Xing X-X. Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: A systems neuroscience perspective. *Neurosci. Biobehav. Rev*. 2014; 45:100–118. DOI: 10.1016/j.neubiorev.2014.05.009 [PubMed: 24875392]





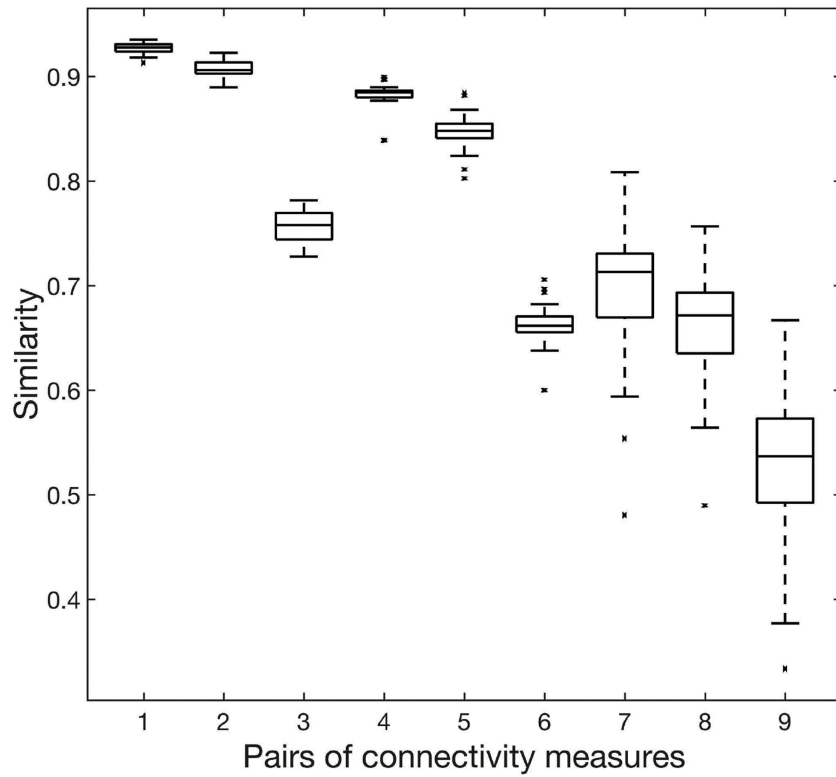
**Figure 1. Performance of 12 Connectome-based Predictive Models (CPMs)**

CPM performance was assessed by correlating predicted and observed individual scores. The twelve models varied according to: definition of functional connectivity (Pearson’s correlation, accordance, or discordance), feature selection and predictive algorithms (linear or PLS regression), and data type on which the model was trained (gradCPT-based fMRI or resting-state fMRI). All predictive models were validated within the gradCPT dataset and applied to three independent datasets (stop-signal task, ANT, and ADHD). Except external dataset 3, which contains only resting-state fMRI, all datasets include both task-based fMRI and resting-state fMRI. Dark red bars and dotted lines represent results from our previously published model (Rosenberg et al., 2016a, b).



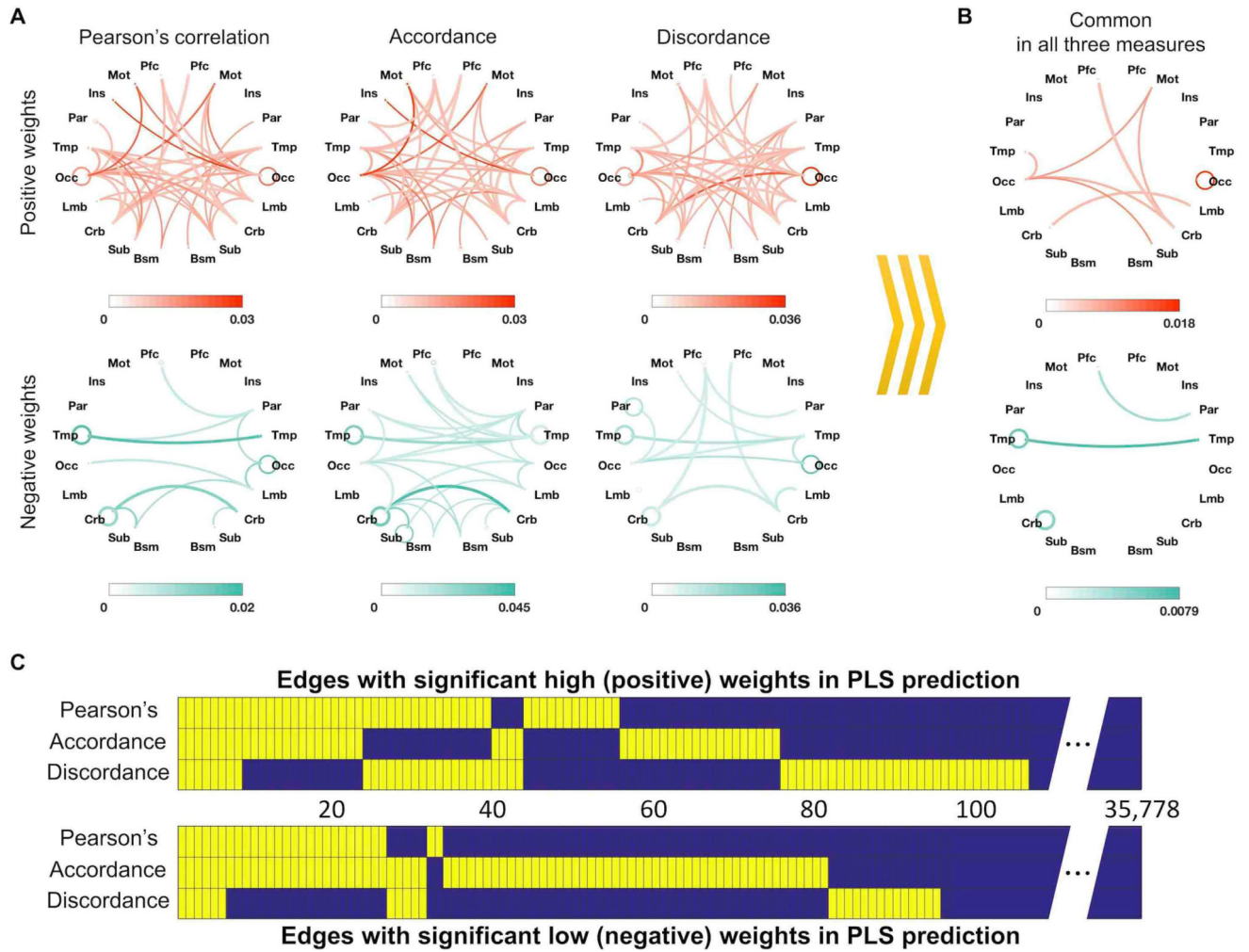
**Figure 2. Average performance of twelve connectome-based predictive models (CPMs) in internal and external datasets**

(A) Mean performance of all models trained on gradCPT task vs. gradCPT rest data. (B) Mean performance of all linear vs. all PLS regression models. (C) Mean performance of all Pearson's correlations vs. accordance vs. discordance models. (D) For the internal validation columns, bars represent predictive power averaged across connectivity measures. For the external validation columns, bars represent predictive power averaged across connectivity measures and datasets. For example, the "task external" group came from averaging the results from testing on stop-signal task-based fMRI and testing on ANT-based fMRI. The "rest external" group was obtained from averaging the results from testing on resting-state fMRI from the three external datasets.



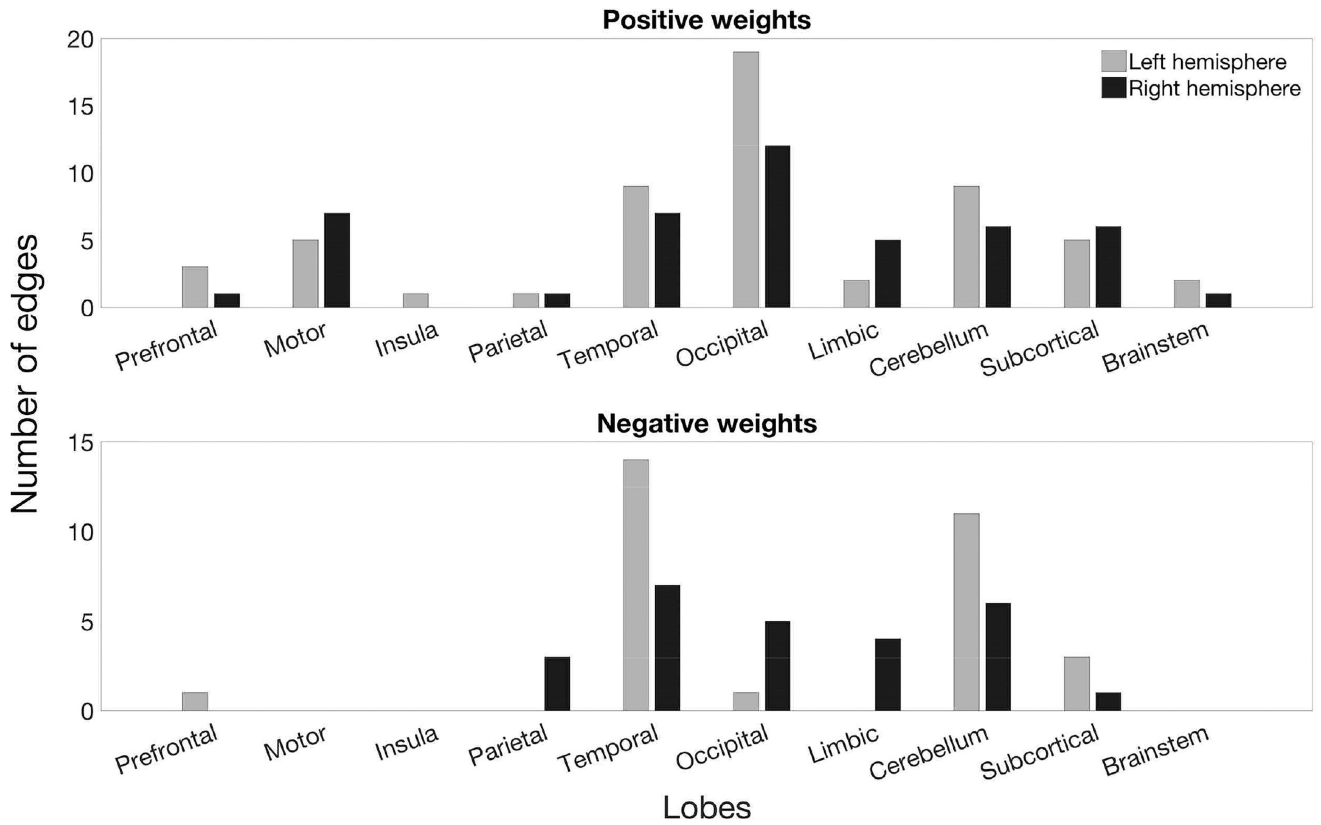
**Figure 3. Similarity among connectivity measures and between sessions**

The first three pairs represent similarities between 1) correlation and accordance, 2) correlation and discordance, and 3) accordance and discordance, all calculated from gradCPT-based fMRI. Pairs 4–6 represent the same pairs as 1–3, but calculated from resting-state fMRI. Pairs 7–9 represent similarities between task fMRI and resting fMRI of 7) correlation, 8) accordance, and 9) discordance.



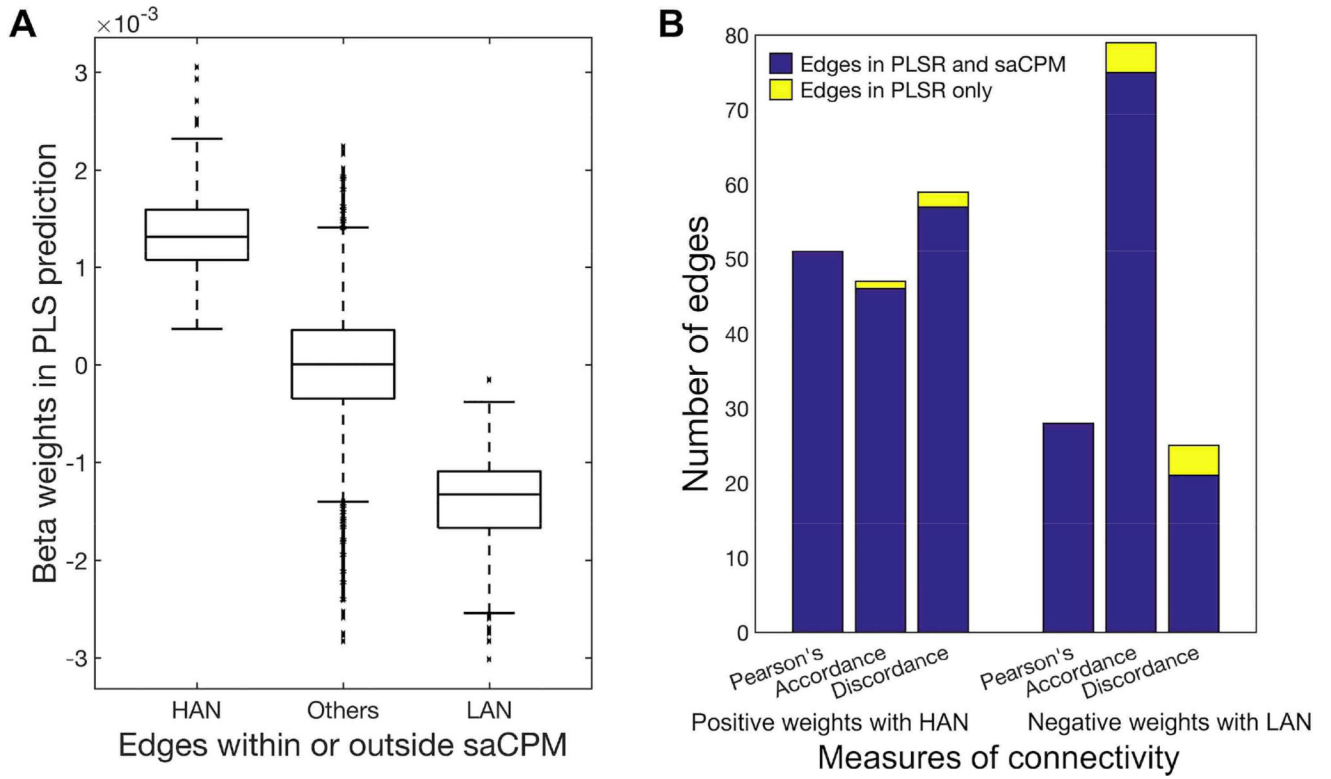
**Figure 3. Edges with significant weights shared across PLS CPMs using different functional connectivity features**

This figure shows edges with a significant beta coefficient in the three PLS prediction models using the three different connectivity measures. **(A)** Anatomical distribution of significant edges in each model; color saturation represents proportion relative to total number of possible edges between each pair of anatomical regions. **(B)** Edges that had significant weights in all three models. **(C)** Significant edges are shown in yellow for each model, and common edges across models are aligned vertically. 106 edges (of a possible 35,778) had significant positive weights, and 95 edges had significant negative weights in at least one connectivity measure (two-tailed  $p < 0.05$  based on 106 permutations).



**Figure 5. Edges with significant weights in the model using PLS regression to predict attention from Pearson’s correlation coefficients**

Bar plots show the number of edges with at least one node in each lobe. Occipital, temporal, and cerebellar regions had the highest number of edges with significant positive weights. Temporal and cerebellar regions also had the highest number of edges with significant negative weights. Every edge is represented twice in this figure – for example, a prefrontal-temporal connection is counted in both the prefrontal and temporal bars.



**Figure 6. Comparison between significant edges from PLS and linear regression**  
**(A)** PLS regression beta weights in the saCPM’s high-attention network (HAN), low-attention network (LAN), and in neither. **(B)** Edges with significant PLS weights that are also included in the HAN or LAN are shown in blue; edges with significant PLS weights that are *not* included in the HAN or LAN are shown in yellow.



Table 1

Prediction performances of 12 connectome-based models in four different datasets

Task	Models trained on resting-state fMRI (from Internal dataset)											
	Models trained on task-based fMRI (from Internal dataset)						Models trained on resting-state fMRI (from Internal dataset)					
	Linear regression			PLS regression			Linear regression			PLS regression		
	Pearson's correlation	Accordance	Discordance	Pearson's correlation	Accordance	Discordance	Pearson's correlation	Accordance	Discordance	Pearson's correlation	Accordance	Discordance
Internal gradCPT	0.875 <sup>*</sup> (1.03×10 <sup>-8</sup> )	0.896 <sup>*</sup> (1.40×10 <sup>-9</sup> )	0.875 <sup>*</sup> (1.02×10 <sup>-8</sup> )	0.884 <sup>*</sup> (4.86×10 <sup>-9</sup> )	0.908 <sup>*</sup> (3.54×10 <sup>-10</sup> )	0.893 <sup>*</sup> (1.93×10 <sup>-9</sup> )	0.738 <sup>*</sup> (2.50×10 <sup>-5</sup> )	0.793 <sup>*</sup> (2.25×10 <sup>-6</sup> )	0.675 <sup>*</sup> (2.17×10 <sup>-4</sup> )	0.790 <sup>*</sup> (2.69×10 <sup>-6</sup> )	0.783 <sup>*</sup> (3.68×10 <sup>-6</sup> )	0.787 <sup>*</sup> (3.01×10 <sup>-6</sup> )
Rest	0.473 <sup>*</sup> (1.70×10 <sup>-2</sup> )	0.567 <sup>*</sup> (3.14×10 <sup>-3</sup> )	0.554 <sup>*</sup> (6.02×10 <sup>-3</sup> )	0.527 <sup>*</sup> (6.75×10 <sup>-3</sup> )	0.510 <sup>*</sup> (9.29×10 <sup>-3</sup> )	0.554 <sup>*</sup> (4.05×10 <sup>-3</sup> )	0.400 <sup>*</sup> (4.73×10 <sup>-2</sup> )	0.530 <sup>*</sup> (6.44×10 <sup>-3</sup> )	0.217 <sup>*</sup> (2.97×10 <sup>-1</sup> )	0.383 <sup>*</sup> (5.89×10 <sup>-2</sup> )	0.372 <sup>*</sup> (6.68×10 <sup>-2</sup> )	0.324 <sup>*</sup> (1.14×10 <sup>-1</sup> )
External 1 SST	0.595 <sup>*</sup> (3.09×10 <sup>-9</sup> )	0.554 <sup>*</sup> (5.65×10 <sup>-8</sup> )	0.490 <sup>*</sup> (2.64×10 <sup>-6</sup> )	0.573 <sup>*</sup> (1.48×10 <sup>-8</sup> )	0.576 <sup>*</sup> (1.22×10 <sup>-8</sup> )	0.566 <sup>*</sup> (2.41×10 <sup>-8</sup> )	0.492 <sup>*</sup> (2.28×10 <sup>-6</sup> )	0.355 <sup>*</sup> (9.95×10 <sup>-4</sup> )	0.370 <sup>*</sup> (5.74×10 <sup>-4</sup> )	0.488 <sup>*</sup> (2.85×10 <sup>-6</sup> )	0.488 <sup>*</sup> (2.87×10 <sup>-6</sup> )	0.458 <sup>*</sup> (1.37×10 <sup>-5</sup> )
Rest	0.294 <sup>*</sup> (1.22×10 <sup>-2</sup> )	0.300 <sup>*</sup> (1.05×10 <sup>-2</sup> )	0.228 <sup>*</sup> (5.42×10 <sup>-2</sup> )	0.233 <sup>*</sup> (4.91×10 <sup>-2</sup> )	0.255 <sup>*</sup> (4.60×10 <sup>-2</sup> )	0.235 <sup>*</sup> (3.63×10 <sup>-2</sup> )	0.322 <sup>*</sup> (5.85×10 <sup>-3</sup> )	0.217 <sup>*</sup> (6.77×10 <sup>-2</sup> )	0.324 <sup>*</sup> (5.47×10 <sup>-3</sup> )	0.395 <sup>*</sup> (5.95×10 <sup>-4</sup> )	0.393 <sup>*</sup> (6.47×10 <sup>-4</sup> )	0.396 <sup>*</sup> (5.85×10 <sup>-4</sup> )
External 2 ANT	-0.654 <sup>*</sup> (3.55×10 <sup>-6</sup> )	-0.637 <sup>*</sup> (7.72×10 <sup>-6</sup> )	-0.590 <sup>*</sup> (4.87×10 <sup>-5</sup> )	-0.631 <sup>*</sup> (9.56×10 <sup>-6</sup> )	-0.629 <sup>*</sup> (1.06×10 <sup>-5</sup> )	-0.623 <sup>*</sup> (1.35×10 <sup>-5</sup> )	-0.595 <sup>*</sup> (4.03×10 <sup>-5</sup> )	-0.517 <sup>*</sup> (5.38×10 <sup>-4</sup> )	-0.547 <sup>*</sup> (2.15×10 <sup>-4</sup> )	-0.593 <sup>*</sup> (4.44×10 <sup>-5</sup> )	-0.581 <sup>*</sup> (6.72×10 <sup>-5</sup> )	-0.581 <sup>*</sup> (6.88×10 <sup>-5</sup> )
Rest	-0.330 <sup>*</sup> (3.52×10 <sup>-2</sup> )	-0.287 <sup>*</sup> (6.93×10 <sup>-2</sup> )	-0.273 <sup>*</sup> (8.48×10 <sup>-2</sup> )	-0.320 <sup>*</sup> (4.13×10 <sup>-2</sup> )	-0.319 <sup>*</sup> (4.23×10 <sup>-2</sup> )	-0.279 <sup>*</sup> (7.75×10 <sup>-2</sup> )	-0.255 <sup>*</sup> (1.08×10 <sup>-1</sup> )	-0.195 <sup>*</sup> (2.21×10 <sup>-1</sup> )	-0.215 <sup>*</sup> (1.78×10 <sup>-1</sup> )	-0.341 <sup>*</sup> (2.92×10 <sup>-2</sup> )	-0.331 <sup>*</sup> (3.44×10 <sup>-2</sup> )	-0.332 <sup>*</sup> (3.42×10 <sup>-2</sup> )
External 3 ADHD	-0.355 <sup>*</sup> (1.16×10 <sup>-4</sup> )	-0.370 <sup>*</sup> (5.56×10 <sup>-5</sup> )	-0.337 <sup>*</sup> (2.60×10 <sup>-4</sup> )	-0.346 <sup>*</sup> (1.71×10 <sup>-4</sup> )	-0.389 <sup>*</sup> (2.06×10 <sup>-5</sup> )	-0.327 <sup>*</sup> (4.13×10 <sup>-4</sup> )	-0.301 <sup>*</sup> (1.20×10 <sup>-3</sup> )	-0.288 <sup>*</sup> (2.01×10 <sup>-3</sup> )	-0.300 <sup>*</sup> (1.23×10 <sup>-3</sup> )	-0.364 <sup>*</sup> (7.25×10 <sup>-5</sup> )	-0.413 <sup>*</sup> (5.44×10 <sup>-6</sup> )	-0.353 <sup>*</sup> (1.25×10 <sup>-4</sup> )

Correlation between predicted and observed scores (corresponding p-value). Correlations between predicted and observed scores in internal and external validations were expected to be positive, whereas correlations in external validation 2 and 3 were to be negative because of the behavioral/clinical scores used.

\* for p-value < 0.05; **BOLD** for p-value < 1×10<sup>-4</sup>