

A generalizable connectome-based marker of in-scan sustained attention in neurodiverse youth

Authors

Corey Horien^{1,2*}, Abigail S. Greene^{1,2}, Xilin Shen³, Diogo Fortes⁴, Emma Brennan-Wydra⁴, Chitra Banarjee⁴, Rachel Foster⁴, Veda Donthireddy⁴, Maureen Butler⁴, Kelly Powell⁴, Angelina Verneti⁴, Francesca Mandino³, David O'Connor⁵, Evelyn M. R. Lake³, James C. McPartland^{4,6}, Fred R. Volkmar^{4,6}, Marvin Chun⁶, Katarzyna Chawarska^{4,7,8}, Monica D. Rosenberg^{9,10}, Dustin Scheinost^{1,3,4,7}, R. Todd Constable^{1,3,11*}

Affiliations

¹Interdepartmental Neuroscience Program, Yale School of Medicine, New Haven, CT, USA

²MD-PhD Program, Yale School of Medicine, New Haven, CT, USA

³Department of Radiology and Biomedical Imaging, Yale School of Medicine, New Haven, CT, USA

⁴Yale Child Study Center, New Haven, CT, USA

⁵Department of Biomedical Engineering, Yale University, New Haven, CT, USA

⁶Department of Psychology, Yale University, New Haven, CT, USA

⁷Department of Statistics and Data Science, Yale University, New Haven, CT, USA

⁸Department of Pediatrics, Yale School of Medicine, New Haven, CT, USA

⁹Department of Psychology, University of Chicago, United States

¹⁰Neuroscience Institute, University of Chicago, United States

¹¹Department of Neurosurgery, Yale School of Medicine, New Haven, CT, USA

*Corresponding authors:

Corey Horien

Magnetic Resonance Research Center

300 Cedar St

PO Box 208043

New Haven, CT 06520-8043

(203) 785-6199

corey.horien@yale.edu

Todd Constable

Magnetic Resonance Research Center

300 Cedar St

PO Box 208043

New Haven, CT 06520-8043

(203) 785-6199

todd.constable@yale.edu

Brief title

A generalizable marker of attention in neurodiverse youth

Abstract

Difficulty with attention is an important symptom in many conditions in psychiatry, including neurodiverse conditions such as autism. There is a need to better understand the neurobiological correlates of attention and leverage these findings for individuals in healthcare settings. Nevertheless, it remains unclear if it is possible to build robust dimensional predictive models of attention in neurodiverse populations. Here, we use five datasets to identify and validate functional connectome-based markers of attention. In dataset one, we use connectome-based predictive modelling and observe successful prediction of performance on an in-scan sustained attention task in a neurodiverse sample of youth. The predictions are not driven by confounds, such as head motion. In dataset two, we find the attention network model defined in dataset one generalizes to predict in-scan attention in a separate sample of neurotypical participants performing the same attention task. In datasets three to five, we use connectome-based identification and longitudinal scans to probe the stability of the attention network across months to years in individual participants. Our results help elucidate the brain correlates of attention in neurodiverse youth and support the further development of predictive dimensional models of other clinically-relevant phenotypes.

Keywords

Fingerprinting, Functional connectivity, Individual differences, Machine learning, Predictive modelling.

Introduction

Autism spectrum disorder (hereafter ‘autism’) affects approximately 1% of children in the world (Zeidan et al. 2022) and is characterized by impairments in social communication and interaction as well as restricted and repetitive behaviors and atypical responses to sensory information (American Psychiatric Association 2013). An important symptom in autism with widespread individual differences is difficulty with attention. Between ~40-80% (Gadow et al. 2006; Lee and Ousley 2006) of individuals with autism have co-occurring attention symptoms, affecting quality of life (Masi et al. 2017). In addition, other neurodiverse individuals, like those with attention-deficit/hyperactivity-disorder (ADHD) and/or the broader autism phenotype (Ingersoll 2010) also have difficulties with attention (Gerdtts and Bernier 2011; American Psychiatric Association 2013). Given the impact, there has been much recent work investigating the neurobiological correlates of attention in neurodiverse populations through the use of fMRI. Of particular interest have been functional connectivity studies, in which measures of synchrony of the blood-oxygen-level-dependent (BOLD) signal are calculated between different regions of interest (Biswal et al. 1995). Group-based functional connectivity studies—comparing those with a neurodiverse condition like autism (Di Martino et al. 2013; Keehn et al. 2013; Fitzgerald et al. 2015) or ADHD (Qiu et al. 2011; Di Martino et al. 2013; Posner et al. 2013; Hoekzema et al. 2014) to neurotypical participants—have helped advance understanding of the brain correlates of attention.

While these studies have proven useful, they have largely failed to make a clinical impact. One potential reason is a lack of prediction-based studies focusing on individuals. Studies using cross-dataset prediction—building and validating models in one sample, then testing it in a separate sample (Scheinost et al. 2019)—are rare in the neuroimaging literature, despite their potential clinical utility (Gabrieli et al. 2015). Aside from clinical applications, a

prediction-based approach holds promise for avoiding statistical issues hindering generalizability (Yarkoni and Westfall 2017; Yarkoni 2020) and avoids the general lack of reliability of simple association studies (Marek et al. 2022). Finally, a prediction-based framework can offer insights into populations undergoing significant developmental changes, particularly in youth (Rosenberg et al. 2018), and aligns with the goals of interrogating symptom dimensions among diverse individuals to aid further understanding of mental disorders (Insel et al. 2010).

Based on the importance of attention and the need for prediction studies focusing on individual differences, we set out to test if it is possible to build predictive models of sustained attention phenotypes based on an in-scan attention task in a sample of youth with autism and other neurodiverse conditions, as well as neurotypical controls. There are numerous reasons it might not be possible to generate a predictive model in a youth sample comprising many patients, including difficulties with task completion and issues with scan compliance (Yerys et al. 2009). In addition, some have suggested that brain differences among those with a neurodiverse condition compared to neurotypical participants (Ross and Margolis 2019) might make it difficult to dimensionally model a phenotype using measures of functional organization. Furthermore, cognitive processes like attention are supported by complex, brain-wide correlates (Kessler et al. 2016; Rosenberg, Finn, et al. 2016). Is it possible to generate a complex, brain-wide model in a sample comprising neurotypical and neurodiverse individuals?

With these factors in mind, we address three main issues in this work. We aim to 1) determine if attention-based predictive models can be generated in a neurodiverse developmental dataset, 2) test if such a model generalizes out of sample, and 3) interrogate neuroanatomy of the network model and assess the stability in individual participants across time. Using connectome-based predictive modelling (CPM) (Finn et al. 2015; Rosenberg, Finn, et al. 2016; Shen et al.

2017; Beaty et al. 2018; Greene et al. 2018; Hsu et al. 2018; Yoo et al. 2018; Rapuano et al. 2020; Rohr et al. 2020; Boyle et al. 2022), we show that we are indeed able to predict performance on an in-scan sustained attention task in novel subjects based on functional connectivity data. The predictions are robust to factors such as in-scanner head motion, Autism Diagnostic Observation Schedule (ADOS) scores, age, sex, and intelligence quotient (IQ) scores. Crucially, we find the network model generalizes out of sample, increasing confidence in the model. In line with other dimensional work in neurodiverse populations (Lake et al. 2019; Rohr et al. 2020; Xiao et al. 2021), we observe the brain correlates identified by the model are complex and distributed across broad swaths of cortical, subcortical, and cerebellar regions. Using connectome-based identification (ID) (Finn et al. 2015; Kaufmann et al. 2017; Vanderwal et al. 2017; Waller et al. 2017; Amico and Goni 2018; Graff et al. 2022; Graff et al. 2022), we perform exploratory analyses testing the longitudinal stability of the network model in individual participants. In sum, our data suggest robust network markers of attention can be generated in neurodiverse youth and add to the growing literature suggesting the power of dimensional approaches in modelling brain-behavior relationships.

Materials and Methods

Description of datasets

We used five independent data sets (Table 1) in this study. The first dataset consisted of youths with autism and other neurodiverse conditions (e.g., ADHD, anxiety, broader autism phenotype, bipolar disorder) as well as typically developing children and has been described previously (hereafter ‘neurodiverse sample’) (Horien et al. 2020). Participants were scanned on a 3T Siemens Prisma system. See Supplement for exclusion criteria and imaging parameters for the neurodiverse sample. A second dataset of neurotypical adults was used as a test dataset

(hereafter ‘test sample’) and is described elsewhere (Rosenberg, Finn, et al. 2016). Participants were scanned on a 3T Siemens Trio TIM system.

Three additional datasets from the Consortium for Reliability and Reproducibility (CoRR) (Zuo et al. 2014) were used to assess stability of the network model: the University of Pittsburgh School of Medicine dataset, the University of Utah dataset, and the University of McGill dataset (hereafter, ‘Pitt’, ‘Utah’, and ‘UM’, respectively). Full details of the Pitt and UM datasets can be found elsewhere (Hwang et al. 2013; Orban et al. 2015). All scans were acquired using Siemens 3-T Tim Trio scanners; all participants were neurotypical.

All datasets were collected in accordance with the institutional review board or research ethics committee at each site. Where appropriate, informed consent was obtained from the parents or guardians of participants. Written assent was obtained from children ages 13–17; verbal assent was obtained from participants under the age of 13.

Measure	Neurodiverse sample	Utah	UM	Pitt	Test sample
Number of participants (males)	70 (39)	16 (16)	27 (4)	44 (21)	25 (12)
Number of participants with a neurodevelopmental or psychiatric condition	33 total 7 = ADHD 2 = anxiety disorder 20 = autism 3 = BAP 1 = bipolar	-	-	-	-

Age in years, mean (standard deviation)	11.59 (2.87)	20.23 (8.28)	65.3□ (6.28)	15.23□(2.83)	22.79 (3.54)
Time between scans in years, mean (standard deviation)	-	2.54□ (0.29)	0.305□ (0.067)	1.76□(0.41)	-
Scan duration in minutes (volumes)	10 min (two 5 min runs; 600 total volumes)	8 min (240)	5 min (150)	5 min (200)	36 min task (three 12 min runs; 824 total volumes); 6 min rest (360 total volumes)
TR in seconds	1	2	2	1.5	1
IQ, mean (standard deviation)	107.23 (15.93)	-	-	-	-

Table 1. Demographic and imaging characteristics of samples used in this study.

ADHD, attention-deficit/hyperactivity disorder; BAP, broader autism phenotype; ADOS, autism diagnostic observations schedule; IQ, intelligence quotient, TR, repetition time.

gradCPT task description

Participants in the neurodiverse sample completed the gradual onset continuous performance task (gradCPT) (Esterman et al. 2013; Rosenberg et al. 2013; Rosenberg, Finn, et al. 2016). The gradCPT task is an assessment of sustained attention and inhibition abilities that has been shown to produce a range of performance scores across neurotypical participants

(Esterman et al. 2013; Rosenberg et al. 2013). In the task, participants viewed grayscale images of city and mountain scenes presented at the center of the screen. In each trial, an image transitioned from one to the next through linear pixel-by-pixel interpolation. Each transition took 1000 ms. For 1000 ms the current scene transitioned from the previous scene, and for the next 1000 ms, it transitioned to the next. Subjects were told to respond by pressing a button for city scenes and to withhold button presses for mountain scenes. City scenes occurred randomly 90% of the time. As in previous studies (Esterman et al. 2013; Rosenberg, Finn, et al. 2016), accuracy was emphasized without reference to speed, and performance was quantified using d' (sensitivity), the participant's hit rate minus false alarm rate. Participant d' scores were calculated for scan 1 and scan 2 individually, as well as the average across both scans. Participants in the test sample also completed gradCPT with the same parameters as above, except scene transitions took 800 ms; resting-state data were also collected in this sample.

The Pitt, Utah, and UM subjects completed only resting-state scans that were spaced apart at longer time intervals (months to years between scans; Table 1).

Preprocessing

The preprocessing strategy for the neurodiverse sample has been described previously (Greene et al. 2018; Horien et al. 2019). Preprocessing steps were performed using BioImage Suite (Joshi et al. 2011) unless otherwise indicated, and included: skull-stripping the 3D magnetization prepared rapid gradient echo (MPRAGE) images using optiBET (Lutkenhoff et al. 2014) and performing linear and non-linear transformations to warp a 268-node functional atlas from MNI space to single subject space (Greene et al. 2018). Functional images were motion corrected using SPM8 (<https://www.fil.ion.ucl.ac.uk/spm/software/spm8/>). Covariates of no interest were regressed from the data, including linear, quadratic, and cubic drift, a 24-parameter

model of motion (Satterthwaite et al. 2013) mean cerebrospinal fluid signal, mean white matter signal, and the global signal. Data were temporally smoothed with a zero-mean unit-variance low-pass Gaussian filter (approximate cutoff frequency of 0.12 Hz). The results of skull-stripping, non-linear, and linear registrations were inspected visually after each step.

We used previously preprocessed data for the Utah, UM, Pitt, and test samples; the preprocessing approach has been described elsewhere (Horien et al. 2019; Rosenberg, Finn, et al. 2016). (See Supplement for more about how motion was controlled in all analyses in all samples.)

Node and network definition

We used a 268-node functional atlas (Finn et al. 2015). For each participant, the mean time-course of each region of interest (“node” in graph theory) was calculated, and the Pearson correlation coefficient was calculated between each pair of nodes to achieve a symmetric 268 x 268 matrix of correlation values representing “edges” (connections between nodes) in graph theoretic terminology. We transformed the Pearson correlation coefficients to z -scores via a Fisher transformation and only considered the upper triangle of the matrix, yielding 35,778 unique edges for whole-brain CPM analyses.

Connectome-based predictive modelling

To predict gradCPT performance (d') from the brain data (connectivity matrices) in the neurodiverse sample, we used CPM (Shen et al. 2017). (See Supplemental Methods and Supplemental Figure 1 for more details about CPM.)

Briefly, using 10-fold cross-validation, connectivity matrices from gradCPT and d' scores were divided into an independent training set including subjects from 9 folds and a testing set including the left-out fold. In the training set, linear regression was used to relate edge strength to

d' . The edges most strongly related to d' were selected (using a feature selection threshold of $P = 0.05$) for both a ‘high network’ (in which increased connectivity was associated with a higher d' score) and a ‘low network’ (in which decreased connectivity was associated with a higher d' score). Mean network strength was calculated in both the high and low networks, and the difference between these network strengths was calculated (‘combined network strength’), as in previous work (Greene et al. 2018):

$$\text{High network strength}_s = \frac{1}{b} \left(\sum_{i,j} c_{i,j} m^+_{i,j} \right) ; b = \frac{i(j-1)}{2}$$

$$\text{Low network strength}_s = \frac{1}{b} \left(\sum_{i,j} c_{i,j} m^-_{i,j} \right) ; b = \frac{i(j-1)}{2}$$

$$\text{Combined network strength}_s = \text{High network strength}_s - \text{low network strength}_s$$

where c is the connectivity matrix for subject s and m^+ and m^- are binary matrices indexing the edges (i, j) that survived the feature selection threshold for the high or low network. (Recall that c and m^+ and m^- comprise only the upper triangle of the connectivity matrix, as specified above.) Throughout the text, we refer to the edges in the high and low network as comprising the ‘attention network.’

A linear model was then generated relating combined network strength to d' scores in the training data. In a final step, combined network strength was calculated for the left-out participants in the testing set, and the model was applied to generate d' predictions for these left-out subjects. We conducted the main CPM analyses by constructing an average connectivity matrix per participant across the two gradCPT runs; behavioral data were averaged as well, as in previous work using gradCPT (Rosenberg, Finn, et al. 2016). (See ‘Multiverse analysis and CPM’ section below for how the effects of arbitrary choices were assessed.)

As in (Scheinost et al. 2021), model performance was assessed by comparing the similarity between predicted and observed gradCPT d' scores using both Spearman’s correlation

(to avoid distribution assumptions) and root mean square error (defined as:

$$\text{RMSE}(\text{predicted}, \text{observed}) = \sqrt{(1/n \sum_{i=1}^n (\text{actual}_i - \text{predicted}_i)^2)}.$$

We performed 1000 iterations of a given CPM analysis and selected the median-performing model; we report this in the main text when discussing model performance. To calculate significance, we randomly shuffled participant labels and attempted to predict gradCPT d' scores. We repeated this 1000 times and calculated the number of times a permuted predictive accuracy was greater than the median of the unpermuted predictions to achieve a non-parametric P -value:

$$P = (\#\{\rho_{\text{null}} \geq \rho_{\text{median}}\}) / 1000$$

where $\#\{\rho_{\text{null}} \geq \rho_{\text{median}}\}$ indicates the number of permuted predictions numerically greater than or equal to the median of the unpermuted predictions (Scheinost et al. 2021).

Multiverse analysis and CPM

To determine if CPM findings were robust, we used a multiverse approach to explore how results were affected by different analytical choices (Steege et al. 2016). The goal of this approach is not to determine what CPM pipeline gives the ‘best’ prediction; rather, it is to gather converging evidence across a range of analytical scenarios to determine how modelling choices affect results. Specifically, we altered the feature selection threshold used to select significant edges ($P = 0.05$, $P = 0.01$, $P = 0.005$, $P = 0.001$); we tested CPM using a combined network model versus testing the high and low networks separately; we tested the effect of partialling out participant age, sex, IQ, Autism Diagnostic Observation Schedule (ADOS) (Lord et al. 2012) calibrated severity score, and head motion in the feature selection step (as described previously) (Scheinost et al. 2021; Dufford et al. 2022); and we also built models using data from gradCPT run 1, gradCPT run 2, and average gradCPT data. To ensure this approach did not result in false

positives, the Benjamini–Hochberg procedure (Benjamini and Hochberg 1995) was applied to control for multiple comparisons (correcting for 12 tests).

Testing generalizability of the attention network

To assess if the network model of attention generalized out of sample, we defined a consensus high attention network and consensus low attention network as edges that appear in at least 6/10 folds in 600/1000 iterations of CPM. This resulted in 922 edges in the high attention network and 896 edges in the low attention network. Using the combined network strength in the consensus networks (as above for CPM), we determined model coefficients across the neurodiverse sample, as in (Rosenberg, Finn, et al. 2016; Ju et al. 2020; Dufford et al. 2022). We then applied the network masks and model coefficients to the test sample to generate d' predictions.

As above for CPM analyses, model performance was assessed by comparing the similarity between predicted and observed gradCPT d' scores using both Spearman's correlation and by calculating RMSE. Non-parametric P -values were computed as for CPM. As above, we used a similar multiverse approach to ensure results were not being driven by subject age, sex, or head motion; the Benjamini–Hochberg (1995) procedure was again used to control for multiple comparisons. To further ensure results were robust, we tested a range of summary networks of varying sizes (i.e., from stringent cases where an edge must appear in 10/10 folds and 1000/1000 iterations, to more liberal thresholds where an edge must appear in 3/10 folds and 300/1000 iterations).

Connectome-based ID

To test the stability of the predictive model over time in a given individual, we used connectome-based ID (Finn et al. 2015) and the Pitt, Utah, and UM samples. (See Supplemental

Figure 2 for a schematic and more about connectome-based ID.) Briefly, after selecting only the edges in the high and low networks (i.e., the same consensus edges used in the cross-dataset test—the 922 and 896 edges in the high and low attention networks, respectively), a database was created consisting of all subjects' matrices from scan 1. In an iterative process, a connectivity matrix from a given subject was then selected from scan 2 and denoted as the target. Pearson correlation coefficients were calculated between the target connectivity matrix and all the matrices in the database. If the highest Pearson correlation coefficient was between the target subject in one session and the same subject in the second session (i.e. within-subject correlation > all other between subject correlations), this was recorded as a correct identification. The process was repeated until identifications had been performed for all subjects and database-target combinations. We averaged both database-target pairs (because these can be reversed) for a dataset to achieve an average ID rate. To calculate P -values, we randomly shuffled subject identities and reperformed ID for 1000 iterations and compared the actual ID rates to this null distribution (Finn et al. 2015; Horien et al. 2018; Horien et al. 2019):

$$P = (\#\{ID_{\text{null}} \geq ID_{\text{actual}}\}) / 1000$$

where $\#\{ID_{\text{null}} \geq ID_{\text{actual}}\}$ indicates the number of permuted ID rates numerically greater than or equal to the actual ID rate obtained using the original data.

We also assessed if connections inside the attention network were more or less stable than connections in the rest of the brain. We generated 1000 summary networks comprising edges outside of the consensus attention network (and the same size as the high and low attention networks). Connectome-based ID was performed using the random networks, and we compared ID results to those obtained using the original consensus attention network. P -values were obtained as follows:

$$P = (\#\{ID_{\text{random}} \geq ID_{\text{actual}}\}) / 1000$$

where $\#\{ID_{\text{random}} \geq ID_{\text{actual}}\}$ indicates the number of random ID rates numerically greater than or equal to the actual ID rate obtained using the original data.

In the Pitt dataset, incomplete scan coverage during the functional runs resulted in 158/922 and 145/896 edges missing in the high and low networks, respectively; we performed ID with the remaining edges.

Code and data availability

Preprocessing was carried out using software freely available here: (<https://medicine.yale.edu/bioimaging/suite/>). CPM code is available here: (<https://github.com/YaleMRRC/CPM>). The parcellation, the attention network models, and the connectome-based ID code are available here: (https://www.nitrc.org/frs/?group_id=51). Data from the longitudinal samples are openly available through CoRR (http://fcon_1000.projects.nitrc.org/indi/CoRR/html/).

Results

Prediction of in-scan attention scores in the neurodiverse sample

In the neurodiverse sample, there were no differences between neurodiverse and neurotypical participants in in-scanner head motion ($t(68) = 0.77, P = 0.4437$) or gradCPT d' ($t(68) = -0.60, P = 0.5487$). Across the sample, we observed that motion and d' were negatively correlated ($r = -0.35, P = 0.0034$); we hence adopted a rigorous motion control strategy for the duration of the analyses to ensure motion was not driving the brain-behavior models (Supplemental Methods).

Next, using CPM, we built a model using within-dataset cross-validation to predict unseen participants' gradCPT d' scores from functional connectivity data. The model

successfully predicted gradCPT d' scores in the neurodiverse dataset (Spearman's $\rho = 0.54$, RMSE = 0.78, $P = 0.0001$; Figure 1A).

To assess the robustness of the d' prediction, we used a multiverse approach to explore how results were affected by different analytical choices (Steege et al. 2016). We stress the point of this approach is not to determine what pipeline gives the 'best' prediction performance; it is instead to gather converging evidence across a range of analytical scenarios to determine the extent to which arbitrary choices affect CPM results.

As the choice of feature selection threshold is arbitrary, we started by testing a range of thresholds, while still controlling for motion. We were able to significantly predict d' scores in all cases (feature selection threshold of 0.01: Spearman's $\rho = 0.50$, RMSE = 0.79, $P = 0.0001$; feature selection threshold of 0.005: Spearman's $\rho = 0.45$, RMSE = 0.80, $P = 0.002$; feature selection threshold of 0.001: Spearman's $\rho = 0.43$, RMSE = 0.81, $P = 0.0001$). Interestingly, all models performed quite well, but when more stringent feature selection thresholds were applied (i.e., fewer edges were included in a model), performance decreased.

We next performed analyses adjusting for participant age, sex, IQ, and ADOS score. Each pipeline demonstrated similar prediction performance of d' (age-adjusted model: Spearman's $\rho = 0.47$, RMSE = 0.80, $P = 0.001$; sex-adjusted model: Spearman's $\rho = 0.52$, RMSE = 0.77, $P = 0.0001$; IQ-adjusted model: Spearman's $\rho = 0.49$, RMSE = 0.96, $P = 0.003$; ADOS-adjusted model: Spearman's $\rho = 0.53$, RMSE = 0.77, $P = 0.0001$). Models were also built for gradCPT scan 1 and gradCPT scan 2 separately (scan 1: Spearman's $\rho = 0.31$, RMSE = 0.85, $P = 0.031$; scan 2: Spearman's $\rho = 0.27$, RMSE = 0.87, $P = 0.073$). Prediction performance dropped in this case, echoing recent results that more data fed into predictive models results in higher accuracies (Taxali et al. 2021).

Finally, the choice to model attention using a ‘combined network’ (used in all analyses above) is also arbitrary. We repeated the CPM prediction of d' scores and tested the high and low networks. We again observed similar d' prediction in both cases (high network: Spearman’s rho = 0.53, RMSE = 0.78, $P = 0.0001$; low network: Spearman’s rho = 0.50, RMSE = 0.79, $P = 0.0001$; Figure 1B-C).

In all, these results suggest attention prediction is robust in this sample and is not driven by potential confounding factors.

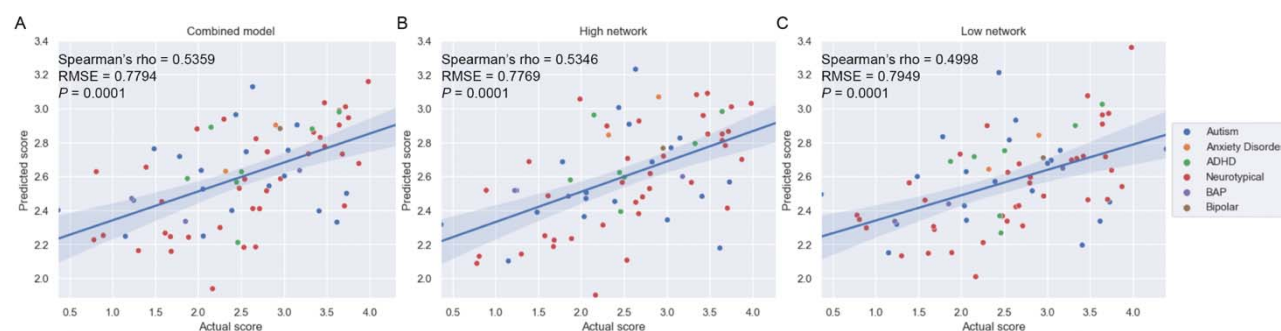


Figure 1. In-scan sustained attention task performance (gradCPT d') can be predicted in a neurodiverse sample using CPM. A) Results from the combined network model. B) Results from the high network. C) Results from the low network. For all plots, actual gradCPT d' scores are indicated on the x-axis; predicted scores, on the y-axis. A regression line and 95% confidence interval are shown. ADHD, attention-deficit/hyperactivity disorder; BAP, broader autism phenotype; $P = P$ -value; RMSE, root mean square error.

External validation of the attention network

Overfitting—deriving statistical patterns specific to noise in a sample—is a constant concern in machine learning studies. The ultimate test is to assess how well a model works in an independent dataset; we perform such a test here. We determined which edges tended to contribute consistently to successful prediction (922 in the high network and 896 in the low

network, 1818 edges total) and built a consensus model in the neurodiverse sample (Methods). Applying the model to the test sample comprising subjects completing the same gradCPT task, we observed successful prediction of d' scores (Spearman's rho = 0.65, RMSE = 1.059 P = 0.0006; Figure 2A).

We repeated analyses controlling for several other variables, adjusting for in-scanner head motion (Spearman's rho = 0.67, P = 0.0008), participant sex (Spearman's rho = 0.61, P = 0.0014), as well as age (Spearman's rho = 0.65, P = 0.0009), and observed similar results. In addition, when we tested both the high and low network separately, we found each of these networks predicted d' scores (high network: Spearman's rho = 0.59, RMSE = 1.094, P = 0.021; low network: Spearman's rho = 0.59, RMSE = 1.03, P = 0.0025). In addition, we tested whether the attention network could be used to predict d' scores from resting-state data in the adult sample. We again found the model generalized (Spearman's rho = 0.42, RMSE = 1.002, P = 0.0358; Figure 2B).

We further tested the stability of results by altering how consistently an edge had to appear across CPM iterations to be included in the summary attention network (Methods). This resulted in 8 attention summary networks, ranging from ~100-3000 edges. In 7/8 cases, the attention network generalized to predict d' scores (range of Spearman's rho = 0.50-0.66; all P < 0.0104 after FDR correction; Supplemental Table 1). The only summary network that did not predict d' score (Spearman's rho = 0.23, P = 0.27) was quite small (~100 edges), approximately an order of magnitude smaller than the original attention network tested above (and other networks that have generalized) (Rosenberg, Finn, et al. 2016; Greene et al. 2018; Yip et al. 2019). These results suggest generalization in this sample is robust to the arbitrary choices made when defining a summary model.

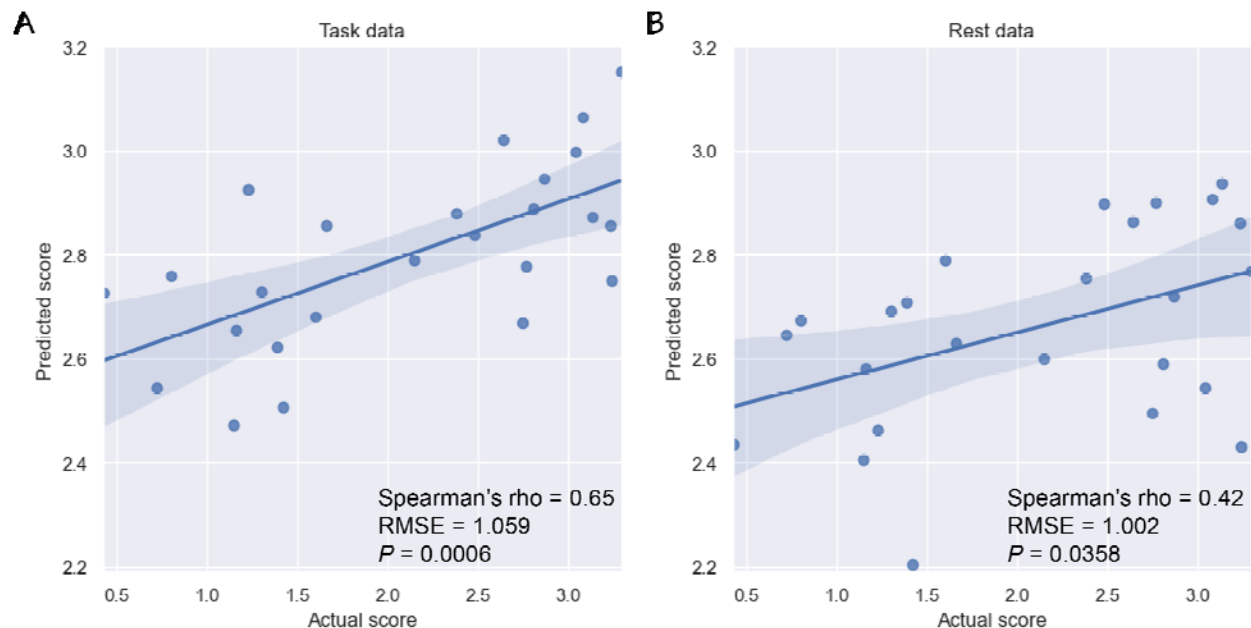


Figure 2. Generalization of the attention network to an independent sample. A) Results using task data. B) Results using rest data. For all plots, actual gradCPT d' scores are indicated on the x-axis; predicted scores, on the y-axis. A regression line and 95% confidence interval are shown. Abbreviations: RMSE, root mean square error; P = P -value.

Neuroanatomy of CPM predictive networks

We next performed post-hoc visualizations to localize brain connections contributing to the model. Together, the 1818 total edges comprise 5.08% of the connectome. Similar to other CPM models (Rosenberg, Finn, et al. 2016; Beaty et al. 2018; Greene et al. 2018; Lake et al. 2019; Ju et al. 2020; Dufford et al. 2022), the predictive edges in the high and low networks comprise complex, distributed networks spanning the entire brain (Figure 3A-B). In line with task demands, additional visualizations at the network level (Figure 3C-D) revealed connections involving subcortical, cerebellar, and visual networks were particularly important. In the high attention network, for example, the top three network pairs containing the greatest proportion of edges involved subcortical, cerebellar, and visual networks. In the low attention network, a

network pair involving the cerebellum (the cerebellar-frontoparietal network) contained the greatest proportion of edges. For completeness, we present the matrices in Figure 3C-D in Supplemental Figure 3 without normalizing by network size; subcortical, visual, and cerebellar networks again tended to harbor large numbers of edges. (See also Supplemental Figure 4 for an additional visualization using circle plots and glass brains.)

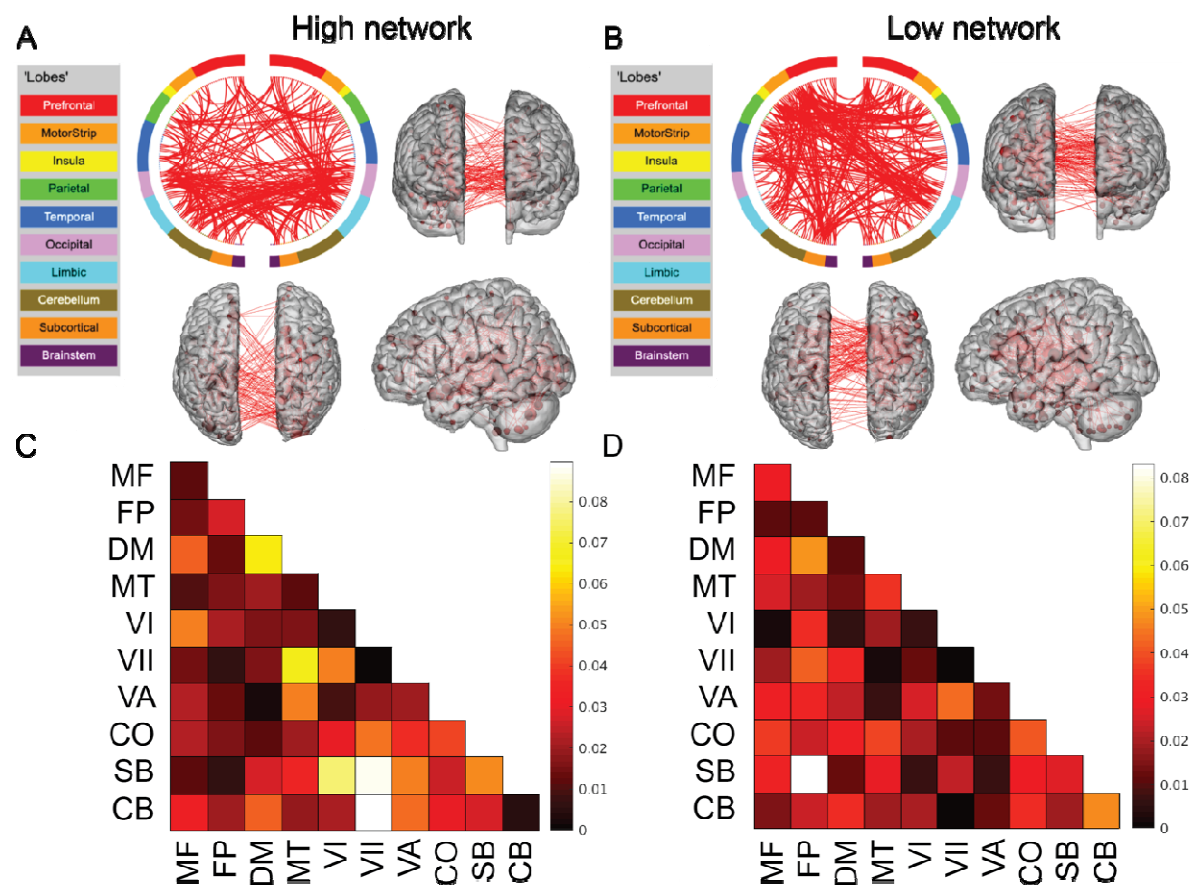


Figure 3. Neuroanatomy of CPM predictive networks. A) The consensus high network. B) The consensus low network. For both A) and B): a circle plot is shown in the upper left. The top of the circle represents anterior; the bottom, posterior. The left half of the circle plot corresponds to the left hemisphere of the brain. A legend indicating the approximate anatomic ‘lobe’ is shown to the left. The same edges are plotted in the glass brains as lines connecting different nodes (red circles); in these visualizations, nodes are sized according to degree, the number of edges

connected to that node. Note that to aid in visualization, we have thresholded the matrices to only show nodes with a degree threshold > 15 (unthresholded circle plots and glass brains are shown in Supplemental Figure 4). C) Matrix of the consensus high network. D) Matrix of the consensus low network. For both C) and D): the proportion of edges in a given network pair; data have been corrected for differing network size. MF, medial frontal; FP, frontoparietal; DM, default mode; MT, motor; VI, visual I; VII, visual II; VA, visual association; CO, cingulo-opercular; SB, subcortical; CB, cerebellum.

Individual-level stability of predictive network model

An ultimate goal of using individual-level approaches in the clinic is to infer future outcomes based on current data. Because the attention network was defined in relation to a state-based cognitive process that itself fluctuates (Cohen and Maunsell 2011; Esterman et al. 2013; Esterman et al. 2014; Rosenberg et al. 2015; Terashima et al. 2021), it is possible individual connectivity patterns in the attention network might change over time. Hence, we conducted an exploratory analysis using connectome-based ID and longitudinal datasets with months to years between scans, asking: are connections in the attention network stable enough within individuals to identify a participant from a group?

Across the three longitudinal samples, we observed that the attention network results in ID rates well-above chance (ID rate range: 53.4 - 93.5%; $P < 0.0001$ across all samples; Figure 4A-B). Specifically, ID rates were high when there were months between scans (UM dataset; 92.6% and 81.5% in the high and low attention networks, respectively) and when there were years between scans (as for Utah and Pitt datasets; Utah high attention network = 84.4%, Utah low attention network = 81.3%; Pitt high attention network = 53.4%, Pitt low attention network

= 62.5%). These results suggest participants retain their unique connectivity signatures in the attention network.

We next examined the relationship between in-scanner head motion and within-participant correlations in connectivity strength in the attention network. Across all three samples, we found there were no statistically significant relationships between within-participant correlation scores and head motion in 5/6 cases (range of rho values: -0.096 - 0.123; $P > 0.53$ across all samples; Supplemental Table 2). The only statistically significant relationship we observed was in Pitt, and there was a negative association (high attention network: rho = -0.36, $P = 0.0153$), indicating higher head motion in this sample was associated with lower within-subject self-correlation scores (in line with previous results) (Horien et al. 2018; Graff et al. 2022). These findings suggest the high ID rates are not being driven by head motion.

Lastly, we used connectome-based ID to determine how the attention network compared to connections in the rest of the brain. Generating 1000 summary networks comprising edges outside of the consensus attention network (and of the same size), we observed no differences in ID rates of the random networks compared to the attention network (ID rate range: 62 - 92.8%; all $P > 0.49$ across all samples; Figure 4C-D; see Supplemental Table 3 for exact ID rates and P -values for all samples). That is, the attention network itself tends to retain participant-specific connectivity signatures across months to years because the brain as a whole retains participant-specific connectivity signatures across months to years.

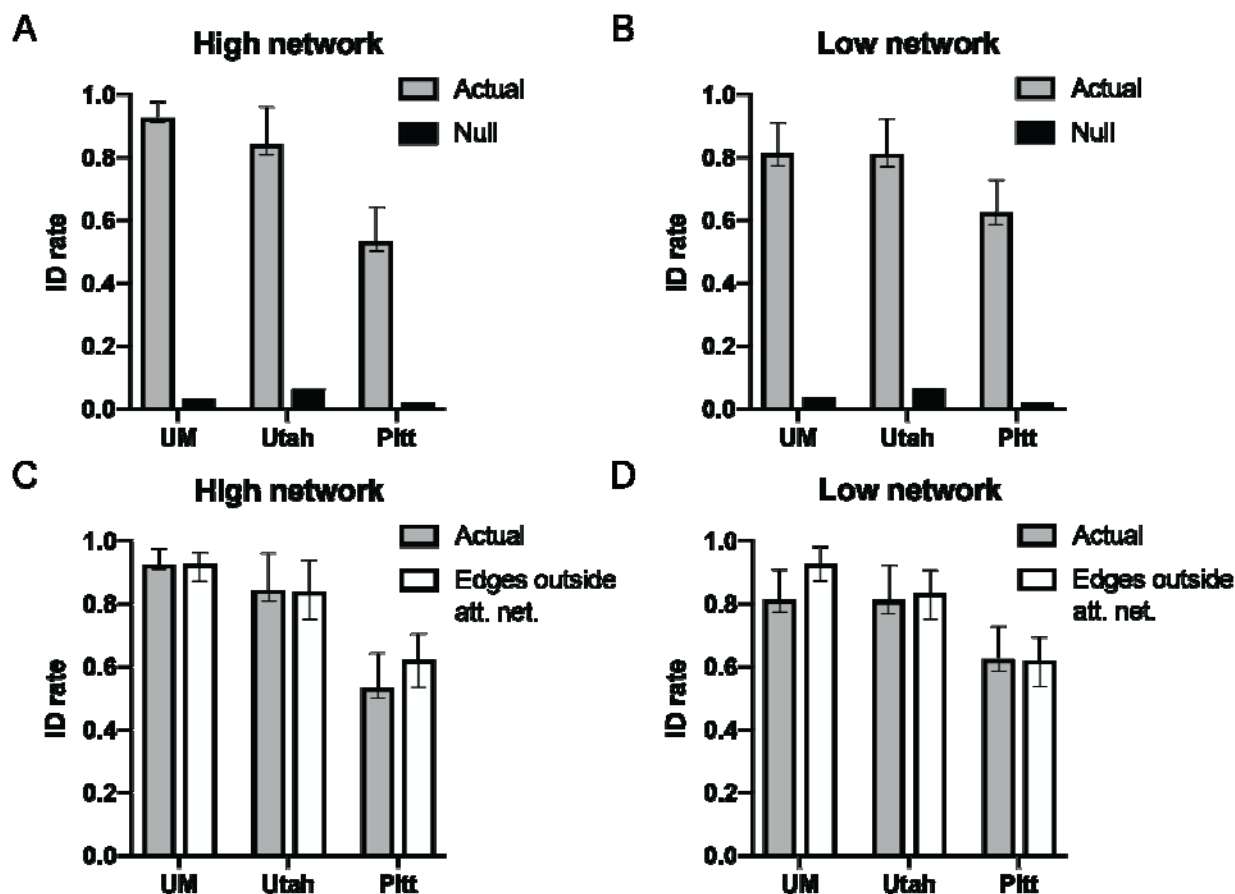


Figure 4. The attention network is stable at the individual level across months and years. For panels A) and B) results are shown for the high and low networks, respectively, comparing ID rates obtained using edges in the attention network ('Actual') compared to chance levels when participant labels have been permuted ('Null'). In A) and B), the rightmost bar is the mean ID rate from 1000 iterations of permutation testing. For panels C) and D) results are shown for the high and low networks, respectively, comparing ID rates obtained using edges in the attention network ('Actual') compared to ID rates obtained when using edges not included in the attention network ('Edges outside att. net.'). For A-D): each dataset is indicated below the x-axis; the y-axis is the ID rate, presented as a fraction from 0 (no subjects correctly identified) to 1 (all subjects correctly identified). Error bars correspond to 95% confidence intervals obtained via

bootstrapping. Note the actual ID rates shown in plots A) and B) (the grey bars) are reproduced in plots C) and D). Att. net., attention network; ID rate, identification rate; Pitt, University of Pittsburgh sample; UM, University of Montreal sample.

Discussion

In this work, we set out to test if it was possible to generate connectome-based predictive models of attention in a neurodiverse sample of youth. Using CPM, we were able to build a predictive network model of in-scan sustained attention scores. Crucially, we found the network generalized out of sample, further suggesting the brain-behavior model we originally identified is a robust marker of attention. The network model was spatially complex, comprising connections across the brain. Lastly, we conducted exploratory analyses in three open-source samples using the network and connectome-based ID.

The power and potential of dimensional models

Our work adds to the growing literature suggesting it is feasible to use a dimensional approach to model individual differences in brain-behavior relationships in neurodiverse youth. In addition to autism symptoms (Lake et al. 2019), groups have developed dimensional models predictive of behavioral inhibition (Rohr et al. 2020), social affect (Xiao et al. 2021), and adaptive functioning (Plitt et al. 2015) in neurodiverse samples. In all cases, the predictive models comprise complex networks, with connections spanning the entire brain (reviewed in) (Horien et al. 2022). Nevertheless, subcortical and cerebellar networks tend to emerge as major contributors in these models, regions we also observed as important in our attention network. Furthermore, similar brain areas have also been noted to play a role in attention (Rosenberg, Finn, et al. 2016; Green et al. 2017; Yoo et al. 2022), consistent with the growing recognition

that subcortical and cerebellar circuits are important in mediating cognitive processes (Buckner 2013; Clark et al. 2021).

Beyond helping to hone in on brain areas involved in sustaining attention, the network identified is intriguing from a clinical standpoint. For example, it has been shown that a network connectivity model of attention (Rosenberg, Finn, et al. 2016) is sensitive to methylphenidate (Rosenberg, Zhang, et al. 2016). It is therefore possible the network identified here may help in tracking changes after administration of a therapeutic. Though more work is needed, it is generally encouraging that markers identified through dimensional analyses appear to be sensitive to clinically useful drugs. There is a need for objective, biological markers in psychiatry, and dimensional approaches could offer a framework to identify quantitative markers to help individuals clinically (McPartland 2021).

Generalizability of the attention network and open science

Much has been written about the reproducibility crisis in biomedical science (Pashler and Wagenmakers 2012; Open Science Collaboration 2015; Baker 2016), as well as the fear results might suffer from a lack of generalizability (Yarkoni 2020), particularly in psychiatry and psychology. We emphasize the present study uses predictive modeling, which is an additional step beyond association studies, that helps reproducibility (Rosenberg, Finn, et al. 2016; Yarkoni and Westfall 2017; Rosenberg et al. 2018; Scheinost et al. 2019; Poldrack et al. 2020). Testing to ensure results are robust across samples and contexts is also imperative. This effort is especially important in neuroimaging, where hopes have been high for clinical impact, yet there has been little progress translating papers into practice (Chekroud 2017; Chekroud and Koutsouleris 2018). Further, even when findings might be clinically useful, many roadblocks stand in the way

of successful implementation (Chekroud and Koutsouleris 2018). It is incumbent on researchers to test findings in multiple samples to avoid having other investigators waste time and resources.

Hence, the fact the attention network generalizes out of sample increases confidence in the original model and opens new opportunities for analysis. That is, the attention network generalized from a young, neurodiverse sample to an older, neurotypical sample. This finding is in line with the dimensional view of brain-behavior relationships (Insel et al. 2010) and also supports the notion that despite developmental changes in brain function, a ‘core’ brain network architecture associated with sustaining attention is likely present (Yoo et al. 2022). An important next step to further assess generalizability will be testing if the network can generalize to predict different aspects of attention (or other phenotypes entirely) and in different populations. To this end, we openly share the attention network model and encourage other investigators to test it widely. By sharing materials, particularly those from neurodiverse participants, we as a community can ensure our findings are relevant for all individuals.

Individual stability of the attention network

Using connectome-based ID, it has been shown the connectome tends to be individually stable across short time scales (minutes) (Miranda-Dominguez et al. 2014; Finn et al. 2015; Kaufmann et al. 2017; Vanderwal et al. 2017; Waller et al. 2017; Amico and Goni 2018) and longer time scales (months to years) (Miranda-Dominguez et al. 2018; Horien et al. 2019; Demeter et al. 2020; Jalbrzikowski et al. 2020; Ousdal et al. 2020; Graff et al. 2022; Graff et al. 2022). Further, individual resting-state networks have tended to exhibit high stability (Finn et al. 2015). Here, we tested if a state-based, behaviorally defined network comprising edges across the brain exhibited the same degree of stability.

Encouragingly, we observed the attention network is a stable multivariate marker of connectivity over long time scales, and this did not differ from the rest of the brain in terms of stability. We interpret this finding as a positive for the field, as it suggests that networks for other phenotypes, with different neurobiological correlates, will likely exhibit a high degree of individual stability as well. Indeed, this seems to be the case, as recent work has demonstrated that predictive network models tend to have substantially higher reliability than individual functional connections (Taxali et al. 2021). Future studies could more rigorously assess how stability of the attention network relates to CPM model performance in longitudinal samples, as the datasets used here did not contain assessments of attention. Continuing to focus on the stability and reliability of our measurements is crucial (Noble et al. 2019), as a lack of reliability can continue to impede the clinical utility of fMRI (Milham et al. 2021)

Limitations and future considerations

Compared to other open source datasets (e.g., ABCD (Casey et al. 2018), the Human Connectome Project (Van Essen et al. 2013), UK Biobank (Miller et al. 2016)), the samples used here are small. Another limitation is participants in the neurodiverse sample had fairly high IQ scores compared to the population at large (Wingate et al. 2014). While IQ was not shown to be a confounding factor in the predictive model, individuals with lower IQs may have difficulties completing the gradCPT task. Future studies could address generalizability of the task and/or network in more varied individuals. In addition, attention is a broad construct, and in this work, we focused on the ability to sustain attention. More research could be conducted to determine if it is possible to build dimensional models of other aspects of attention in neurodiverse individuals.

The longitudinal samples we used to measure stability of the model did not contain behavioral/clinical data. It is unclear if the predictive network is able to predict attention phenotypes across longer time scales. Studies with the same participants completing the gradCPT task at multiple time points could help to answer this question. Finally, we have focused on a single phenotype here. An important next step will be to use a multimodal, multidimensional framework in large numbers of individuals—incorporating numerous phenotypes and data types—to generate findings across multiple spatial and temporal scales (Lombardo et al. 2019). Such an approach holds the promise of illuminating the complex biology underlying heterogeneous neurodiverse conditions.

Conclusion

In sum, we have shown that it is possible to generate task-based predictive models of in-scan attention in a neurodiverse sample and that such a model generalizes. Results support the further development of predictive dimensional models of cognitive phenotypes and suggest that such an approach can yield stable imaging markers.

Acknowledgements and Disclosures

The authors thank Hedwig Sarofin and Cheryl McMurray for technical assistance during the MRI scans and Jitendra Bhawnani for technical assistance with task hardware and software. This work was supported by the National Institutes of Health (P50MH115716, T32GM007205 to CH and ASG, and TR001864 to ASG). JCM consults with Customer Value Partners, Bridgebio, Determined Health, and BlackThorn Therapeutics, has received research funding from Janssen Research and Development, serves on the Scientific Advisory Boards of Pastorius and Modern Clinics, and receives royalties from Guilford Press, Lambert, and Springer.

References

Amico E, Goni J. 2018. The quest for identifiability in human functional connectomes. *Sci Rep.* 8:8254.

American Psychiatric Association . 2013. Diagnostic and statistical manual of mental disorders: DSM-5. Arlington: American Psychiatric Association.

Baker M. 2016. 1,500 scientists lift the lid on reproducibility. *Nature.* 533:452-454.

Beaty RE, Kenett YN, Christensen AP, Rosenberg MD, Benedek M, Chen Q, Fink A, Qiu J, Kwapil TR, Kane MJ, Silvia PJ. 2018. Robust prediction of individual creative ability from brain functional connectivity. *Proc Natl Acad Sci U S A.* 115:1087-1092.

Benjamini Y, Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B.* 57:289-300.

Biswal B, Yetkin FZ, Haughton VM, Hyde JS. 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med.* 34:537-541.

Boyle R, Connaughton M, McGlinchey E, Knight SP, De Looze C, Carey D, Stern Y, Robertson IH, Kenny RA, Whelan R. 2022. Connectome-based predictive modeling of cognitive reserve using task-based functional connectivity. *bioRxiv.2022.2006.2001.494342.*

Buckner RL. 2013. The cerebellum and cognitive function: 25 years of insight from anatomy and neuroimaging. *Neuron.* 80:807-815.

Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM, Soules ME, Teslovich T, Dellarco DV, Garavan H, Orr CA, Wager TD, Banich MT, Speer NK, Sutherland MT, Riedel MC, Dick AS, Bjork JM, Thomas KM, Chaarani B, Mejia MH, Hagler DJ, Jr., Daniela Cornejo M, Sicut CS, Harms MP, Dosenbach NUF, Rosenberg M, Earl E, Bartsch H, Watts R, Polimeni JR, Kuperman JM, Fair DA, Dale AM, Workgroup AIA. 2018. The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev Cogn Neurosci.* 32:43-54.

Chekroud AM. 2017. Bigger Data, Harder Questions-Opportunities Throughout Mental Health Care. *Jama Psychiat.* 74:1183-1184.

Chekroud AM, Koutsouleris N. 2018. The perilous path from publication to practice. *Mol Psychiatr.* 23:24-25.

Clark SV, Semmel ES, Aleksonis HA, Steinberg SN, King TZ. 2021. Cerebellar-Subcortical-Cortical Systems as Modulators of Cognitive Functions. *Neuropsychol Rev.* 31:422-446.

Cohen MR, Maunsell JH. 2011. When attention wanders: how uncontrolled fluctuations in attention affect performance. *J Neurosci.* 31:15802-15806.

Demeter DV, Engelhardt LE, Mallett R, Gordon EM, Nugiel T, Harden KP, Tucker-Drob EM, Lewis-Peacock JA, Church JA. 2020. Functional Connectivity Fingerprints at Rest Are Similar across Youths and Adults and Vary with Genetic Similarity. *Iscience*. 23.

Di Martino A, Zuo XN, Kelly C, Grzadzinski R, Mennes M, Schvarcz A, Rodman J, Lord C, Castellanos FX, Milham MP. 2013. Shared and distinct intrinsic functional network centrality in autism and attention-deficit/hyperactivity disorder. *Biol Psychiatry*. 74:623-632.

Dufford AJ, Kimble V, Tejavibulya L, Dadashkarimi J, Ibrahim K, Sukhodolsky DG, Scheinost D. 2022. Predicting Transdiagnostic Social Impairments in Childhood using Connectome-based Predictive Modeling. *medRxiv.2022.2004.2007.22273518*.

Esterman M, Noonan SK, Rosenberg M, DeGutis J. 2013. In the Zone or Zoning Out? Tracking Behavioral and Neural Fluctuations During Sustained Attention. *Cereb Cortex*. 23:2712-2723.
Esterman M, Rosenberg MD, Noonan SK. 2014. Intrinsic fluctuations in sustained attention and distractor processing. *J Neurosci*. 34:1724-1730.

Finn ES, Shen XL, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT. 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat Neurosci*. 18:1664-1671.

Fitzgerald J, Johnson K, Kehoe E, Bokde AL, Garavan H, Gallagher L, McGrath J. 2015. Disrupted functional connectivity in dorsal and ventral attention networks during attention orienting in autism spectrum disorders. *Autism Res*. 8:136-152.

Gabrieli JDE, Ghosh SS, Whitfield-Gabrieli S. 2015. Prediction as a Humanitarian and Pragmatic Contribution from Human Cognitive Neuroscience. *Neuron*. 85:11-26.

Gadow KD, DeVincent CJ, Pomeroy J. 2006. ADHD symptom subtypes in children with pervasive developmental disorder. *J Autism Dev Disord*. 36:271-283.

Gerds J, Bernier R. 2011. The broader autism phenotype and its implications on the etiology and treatment of autism spectrum disorders. *Autism Res Treat*. 2011:545901.

Graff K, Tansey R, Ip A, Rohr C, Dimond D, Dewey D, Bray S. 2022. Benchmarking common preprocessing strategies in early childhood functional connectivity and intersubject correlation fMRI. *Dev Cogn Neurosci*. 54:101087.

Graff K, Tansey R, Rai S, Ip A, Rohr C, Dimond D, Dewey D, Bray S. 2022. Functional connectomes become more longitudinally self-stable, but not more distinct from others, across early childhood. *Neuroimage*. 258:119367.

Green JJ, Boehler CN, Roberts KC, Chen LC, Krebs RM, Song AW, Woldorff MG. 2017. Cortical and Subcortical Coordination of Visual Spatial Attention Revealed by Simultaneous EEG-fMRI Recording. *J Neurosci*. 37:7803-7810.

Greene AS, Gao SY, Scheinost D, Constable RT. 2018. Task-induced brain state manipulation improves prediction of individual traits. *Nat Commun.* 9.

Hoekzema E, Carmona S, Ramos-Quiroga JA, Richarte Fernandez V, Bosch R, Soliva JC, Rovira M, Bulbena A, Tobena A, Casas M, Vilarroya O. 2014. An independent components and functional connectivity analysis of resting state fMRI data points to neural network dysregulation in adult ADHD. *Hum Brain Mapp.* 35:1261-1272.

Horien C, Floris DL, Greene AS, Noble S, Rolison M, Tejavibulya L, O'Connor D, McPartland JC, Scheinost D, Chawarska K, Lake EMR, Constable RT. 2022. Functional Connectome-Based Predictive Modeling in Autism. *Biol Psychiatry.*

Horien C, Fontenelle S, Joseph K, Powell N, Nutor C, Fortes D, Butler M, Powell K, Macris D, Lee K, Greene AS, McPartland JC, Volkmar FR, Scheinost D, Chawarska K, Constable RT. 2020. Low-motion fMRI data can be obtained in pediatric participants undergoing a 60-minute scan protocol. *Sci Rep-Uk.* 10.

Horien C, Noble S, Finn ES, Shen X, Scheinost D, Constable RT. 2018. Considering factors affecting the connectome-based identification process: Comment on Waller et al. *Neuroimage.* 169:172-175.

Horien C, Shen XL, Scheinost D, Constable RT. 2019. The individual functional connectome is unique and stable over months to years. *Neuroimage.* 189:676-687.

Hsu WT, Rosenberg MD, Scheinost D, Constable RT, Chun MM. 2018. Resting-state functional connectivity predicts neuroticism and extraversion in novel individuals. *Soc Cogn Affect Neurosci.* 13:224-232.

Hwang K, Hallquist MN, Luna B. 2013. The Development of Hub Architecture in the Human Functional Brain Network. *Cereb Cortex.* 23:2380-2393.

Ingersoll B. 2010. Broader autism phenotype and nonverbal sensitivity: evidence for an association in the general population. *J Autism Dev Disord.* 40:590-598.

Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, Sanislow C, Wang P. 2010. Research Domain Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders. *Am J Psychiat.* 167:748-751.

Jalbrzikowski M, Liu F, Foran W, Klei L, Calabro FJ, Roeder K, Devlin B, Luna B. 2020. Functional connectome fingerprinting accuracy in youths and adults is similar when examined on the same day and 1.5-years apart. *Hum Brain Mapp.* 41:4187-4199.

Joshi A, Scheinost D, Okuda H, Belhachemi D, Murphy I, Staib LH, Papademetris X. 2011. Unified framework for development, deployment and robust testing of neuroimaging algorithms. *Neuroinformatics.* 9:69-84.

Ju Y, Horien C, Chen W, Guo W, Lu X, Sun J, Dong Q, Liu B, Liu J, Yan D, Wang M, Zhang L, Guo H, Zhao F, Zhang Y, Shen X, Constable RT, Li L. 2020. Connectome-based models can predict early symptom improvement in major depressive disorder. *J Affect Disord.* 273:442-452.

Kaufmann T, Alnaes D, Doan NT, Brandt CL, Andreassen OA, Westlye LT. 2017. Delayed stabilization and individualization in connectome development are related to psychiatric disorders. *Nat Neurosci.* 20:513-+.

Keehn B, Shih P, Brenner LA, Townsend J, Muller RA. 2013. Functional connectivity for an "island of sparing" in autism spectrum disorder: an fMRI study of visual search. *Hum Brain Mapp.* 34:2524-2537.

Kessler D, Angstadt M, Sripada C. 2016. Growth Charting of Brain Connectivity Networks and the Identification of Attention Impairment in Youth. *Jama Psychiat.* 73:481-489.

Lake EMR, Finn ES, Noble SM, Vanderwal T, Shen X, Rosenberg MD, Spann MN, Chun MM, Scheinost D, Constable RT. 2019. The Functional Brain Organization of an Individual Allows Prediction of Measures of Social Abilities Transdiagnostically in Autism and Attention-Deficit/Hyperactivity Disorder. *Biol Psychiatry.* 86:315-326.

Lee DO, Ousley OY. 2006. Attention-deficit hyperactivity disorder symptoms in a clinic sample of children and adolescents with pervasive developmental disorders. *J Child Adol Psychop.* 16:737-746.

Lombardo MV, Lai MC, Baron-Cohen S. 2019. Big data approaches to decomposing heterogeneity across the autism spectrum. *Mol Psychiatr.* 24:1435-1450.

Lord C RM, DiLavore PC, Risi S, Gotham K, Bishop S. 2012. *Autism Diagnostic Observation Schedule, Second Edition.* Torrance, CA: Western Psychological Services.

Lutkenhoff ES, Rosenberg M, Chiang J, Zhang K, Pickard JD, Owen AM, Monti MM. 2014. Optimized brain extraction for pathological brains (optiBET). *Plos One.* 9:e115551.

Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, Donohue MR, Foran W, Miller RL, Hendrickson TJ, Malone SM, Kandala S, Feczko E, Miranda-Dominguez O, Graham AM, Earl EA, Perrone AJ, Cordova M, Doyle O, Moore LA, Conan GM, Uriarte J, Snider K, Lynch BJ, Wilgenbusch JC, Pengo T, Tam A, Chen J, Newbold DJ, Zheng A, Seider NA, Van AN, Metoki A, Chauvin RJ, Laumann TO, Greene DJ, Petersen SE, Garavan H, Thompson WK, Nichols TE, Yeo BTT, Barch DM, Luna B, Fair DA, Dosenbach NUF. 2022. Reproducible brain-wide association studies require thousands of individuals. *Nature.* 603:654-660.

Masi A, DeMayo MM, Glozier N, Guastella AJ. 2017. An Overview of Autism Spectrum Disorder, Heterogeneity and Treatment Options. *Neurosci Bull.* 33:183-193.

McPartland JC. 2021. Refining biomarker evaluation in ASD. *Eur Neuropsychopharmacol.*

Milham MP, Vogelstein J, Xu T. 2021. Removing the Reliability Bottleneck in Functional Magnetic Resonance Imaging Research to Achieve Clinical Utility. *Jama Psychiat*.

Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, Bartsch AJ, Jbabdi S, Sotiropoulos SN, Andersson JL, Griffanti L, Douaud G, Okell TW, Weale P, Dragonu I, Garratt S, Hudson S, Collins R, Jenkinson M, Matthews PM, Smith SM. 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci*. 19:1523-1536.

Miranda-Dominguez O, Mills BD, Carpenter SD, Grant KA, Kroenke CD, J.T. Nigg JT, Fair DA. 2014. Connectotyping: model based fingerprinting of the functional connectome. *PLoS One*, 9 e111048

Miranda-Dominguez O, Feczko E, Grayson DS, Walum H, Nigg JT, Fair DA. 2018. Heritability of the human connectome: A connectotyping study. *Netw Neurosci*. 2:175-199.

Noble S, Scheinost D, Constable RT. 2019. A decade of test-retest reliability of functional connectivity: A systematic review and meta-analysis. *Neuroimage*. 203:116157.

Open Science Collaboration. 2015. PSYCHOLOGY. Estimating the reproducibility of psychological science. *Science*. 349:aac4716.

Orban P, Madjar C, Savard M, Dansereau C, Tam A, Das S, Evans AC, Rosa-Neto P, Breitner JCS, Bellec P, Grp P-AR. 2015. Test-retest resting-state fMRI in healthy elderly persons with a family history of Alzheimer's disease. *Sci Data*. 2.

Ousdal OT, Kaufmann T, Kolskar K, Vik A, Wehling E, Lundervold AJ, Lundervold A, Westlye LT. 2020. Longitudinal stability of the brain functional connectome is associated with episodic memory performance in aging. *Hum Brain Mapp*. 41:697-709.

Pashler H, Wagenmakers EJ. 2012. Editors' Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence? *Perspect Psychol Sci*. 7:528-530.

Plitt M, Barnes KA, Wallace GL, Kenworthy L, Martin A. 2015. Resting-state functional connectivity predicts longitudinal change in autistic traits and adaptive functioning in autism. *P Natl Acad Sci USA*. 112:E6699-E6706.

Poldrack RA, Huckins G, Varoquaux G. 2020. Establishment of Best Practices for Evidence for Prediction A Review. *Jama Psychiat*. 77:534-540.

Posner J, Rauh V, Gruber A, Gat I, Wang Z, Peterson BS. 2013. Dissociable attentional and affective circuits in medication-naïve children with attention-deficit/hyperactivity disorder. *Psychiatry Res*. 213:24-30.

Qiu MG, Ye Z, Li QY, Liu GJ, Xie B, Wang J. 2011. Changes of brain structure and function in ADHD children. *Brain Topogr.* 24:243-252.

Rapuano KM, Rosenberg MD, Maza MT, Dennis NJ, Dorji M, Greene AS, Horien C, Scheinost D, Todd Constable R, Casey BJ. 2020. Behavioral and brain signatures of substance use vulnerability in childhood. *Dev Cogn Neurosci.* 46:100878.

Rohr CS, Kamal S, Bray S. 2020. Building functional connectivity neuromarkers of behavioral self-regulation across children with and without Autism Spectrum Disorder. *Dev Cogn Neurosci.* 41.

Rosenberg M, Noonan S, DeGutis J, Esterman M. 2013. Sustaining visual attention in the face of distraction: a novel gradual-onset continuous performance task. *Atten Percept Psycho.* 75:426-439.

Rosenberg MD, Casey BJ, Holmes AJ. 2018. Prediction complements explanation in understanding the developing brain. *Nat Commun.* 9.

Rosenberg MD, Finn ES, Constable RT, Chun MM. 2015. Predicting moment-to-moment attentional state. *Neuroimage.* 114:249-256.

Rosenberg MD, Finn ES, Scheinost D, Papademetris X, Shen XL, Constable RT, Chun MM. 2016. A neuromarker of sustained attention from whole-brain functional connectivity. *Nat Neurosci.* 19:165-+.

Rosenberg MD, Zhang S, Hsu WT, Scheinost D, Finn ES, Shen XL, Constable RT, Li CSR, Chun MM. 2016. Methylphenidate Modulates Functional Network Connectivity to Enhance Attention. *J Neurosci.* 36:9547-9557.

Ross CA, Margolis RL. 2019. Research Domain Criteria: Strengths, Weaknesses, and Potential Alternatives for Future Psychiatric Research. *Mol Neuropsychiatry.* 5:218-236.

Satterthwaite TD, Elliott MA, Gerraty RT, Ruparel K, Loughead J, Calkins ME, Eickhoff SB, Hakonarson H, Gur RC, Gur RE, Wolf DH. 2013. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage.* 64:240-256.

Scheinost D, Dadashkarimi J, Finn ES, Wambach CG, MacGillivray C, Roule AL, Niendam TA, Pine DS, Brotman MA, Leibenluft E, Tseng WL. 2021. Functional connectivity during frustration: a preliminary study of predictive modeling of irritability in youth. *Neuropsychopharmacology.* 46:1300-1306.

Scheinost D, Noble S, Horien C, Greene AS, Lake EMR, Salehi M, Gao SY, Shen XL, O'Connor D, Barron DS, Yip SW, Rosenberg MD, Constable RT. 2019. Ten simple rules for predictive modeling of individual differences in neuroimaging. *Neuroimage.* 193:35-45.

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

Steege S, Tuerlinckx F, Gelman A, Vanpaemel W. 2016. Increasing Transparency Through a Multiverse Analysis. *Perspect Psychol Sci.* 11:702-712.

Taxali A, Angstadt M, Rutherford S, Sripada C. 2021. Boost in Test-Retest Reliability in Resting State fMRI with Predictive Modeling. *Cereb Cortex.* 31:2822-2833.

Terashima H, Kihara K, Kawahara JI, Kondo HM. 2021. Common principles underlie the fluctuation of auditory and visual sustained attention. *Q J Exp Psychol (Hove).* 74:705-715.

Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, Consortium WU-MH. 2013. The WU-Minn Human Connectome Project: an overview. *Neuroimage.* 80:62-79.

Vanderwal T, Eilbott J, Finn ES, Craddock RC, Turnbull A, Castellanos FX. 2017. Individual differences in functional connectivity during naturalistic viewing conditions. *Neuroimage.* 157:521-530.

Waller L, Walter H, Kruschwitz JD, Reuter L, Muller S, Erk S, Veer IM. 2017. Evaluating the replicability, specificity, and generalizability of connectome fingerprints. *Neuroimage.* 158:371-377.

Wingate M, Kirby RS, Pettygrove S, Cunniff C, Schulz E, Ghosh T, Robinson C, Lee LC, Landa R, Constantino J, Fitzgerald R, Zahorodny W, Daniels J, Nicholas J, Charles J, McMahon W, Bilder D, Durkin M, Baio J, Christensen D, Van K, Braun N, Clayton H, Goodman A, Doernberg N, Yeargin-Allsopp M, Monitoring ADD. 2014. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010. *Mmwr Surveill Summ.* 63.

Xiao J, Chen H, Shan X, He C, Li Y, Guo X, Chen H, Liao W, Uddin LQ, Duan X. 2021. Linked Social-Communication Dimensions and Connectivity in Functional Brain Networks in Autism Spectrum Disorder. *Cereb Cortex.* 31:3899-3910.

Yarkoni T. 2020. The generalizability crisis. *Behav Brain Sci.* 1-37.

Yarkoni T, Westfall J. 2017. Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspect Psychol Sci.* 12:1100-1122.

Yerys BE, Jankowski KF, Shook D, Rosenberger LR, Barnes KA, Berl MM, Ritzl EK, Vanmeter J, Vaidya CJ, Gaillard WD. 2009. The fMRI success rate of children and adolescents: typical development, epilepsy, attention deficit/hyperactivity disorder, and autism spectrum disorders. *Hum Brain Mapp.* 30:3426-3435.

Yip SW, Scheinost D, Potenza MN, Carroll KM. 2019. Connectome-Based Prediction of Cocaine Abstinence. *Am J Psychiatry*. 176:156-164.

Yoo K, Rosenberg MD, Hsu WT, Zhang S, Li CR, Scheinost D, Constable RT, Chun MM. 2018. Connectome-based predictive modeling of attention: Comparing different functional connectivity features and prediction methods across datasets. *Neuroimage*. 167:11-22.

Yoo K, Rosenberg MD, Kwon YH, Lin Q, Avery EW, Scheinost D, Constable RT, Chun MM. 2022. A brain-based general measure of attention. *Nat Hum Behav*. 6:782-795.

Zeidan J, Fombonne E, Scora J, Ibrahim A, Durkin MS, Saxena S, Yusuf A, Shih A, Elsabbagh M. 2022. Global prevalence of autism: A systematic review update. *Autism Res*. 15:778-790.

Zuo XN, Anderson JS, Bellec P, Birn RM, Biswal BB, Blautzik J, Breitner JCS, Buckner RL, Calhoun VD, Castellanos FX, Chen AT, Chen B, Chen JT, Chen X, Colcombe SJ, Courtney W, Craddock RC, Di Martino A, Dong HM, Fu XL, Gong QY, Gorgolewski KJ, Han Y, He Y, He Y, Ho E, Holmes A, Hou XH, Huckins J, Jiang TZ, Jiang Y, Kelley W, Kelly C, King M, LaConte SM, Lainhart JE, Lei X, Li HJ, Li KM, Li KC, Lin QX, Liu DQ, Liu J, Liu X, Liu YJ, Le GM, Lu J, Luna B, Luo J, Lurie D, Mao Y, Margulies DS, Mayer AR, Meindl T, Meyerand ME, Nan WZ, Nielsen JA, O'Connor D, Paulsen D, Prabhakaran V, Qi ZG, Qiu J, Shao CH, Shehzad Z, Tang WJ, Villringer A, Wang HL, Wang K, Wei DT, Wei GX, Weng XC, Wu XH, Xu T, Yang N, Yang Z, Zang YF, Zhang L, Zhang QL, Zhang Z, Zhang ZQ, Zhao K, Zhen ZL, Zhou Y, Zhu XT, Milham MP. 2014. An open science resource for establishing reliability and reproducibility in functional connectomics. *Sci Data*. 1.