nature portfolio

Corresponding author(s): Corey Horien

Last updated by author(s): Feb 5, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Preprocessing was carried out using freely available software: (https://medicine.yale.edu/bioimaging/suite/; this leads to the main page of BioImage Suite Web 1.2.0 (current build= 1.2.0, 2020/08/25, 11:43:02).
Data analysis	Template scripts that were used for preprocessing of the functional data, as well as CPM code, are available here: (https://github.com/clhorien/tasks_versus_rest_in_autism_prediction/tree/main). CPM analyses were conducted in MATLAB (R2021b).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

ABIDE data are available here: (https://fcon_1000.projects.nitrc.org/indi/abide/). HBN data are available here: (https://fcon_1000.projects.nitrc.org/indi/

cmi_healthy_brain_network/). The functional parcellation used in analysis is available here: (https://www.nitrc.org/frs/?group_id=51). To protect the sensitive nature of data in the adult attention and Yale youth sample, please reach out to the corresponding authors for questions about, or access to, these samples.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	We have reported sex data for all datasets in the text (and include this information in the population characteristics section below). Sex was based on participant self-report. We do not consider gender in the current work. We did not pursue sex-based analyses in the current paper, as all samples are quite small and focusing on males or females only would cut somewhat small samples in half. Nevertheless, we did try incorporate sex into our analyses as appropriate. For instance, we co-varied for sex when building predictive models and when assessing generalizability.
Population characteristics	The first dataset, the Yale youth sample, comprised 63 subjects from a sample described previously (mean age = 11.7 years, st. dev. = 2.8 years; 29 females; mean IQ = 107.8, st. dev. = 15.1). Twenty of the participants had autism; 11 other participants had a neurodevelopmental condition (five with ADHD, two with anxiety disorder, and four were classified as belonging to the broader autism phenotype. Autism symptoms were scored using the Autism Diagnostic Observation Schedule-2 (ADOS-2) and were ascertained by trained clinical psychologists; calibrated severity scores were used in the present work for the social affect subscale (mean = 3.2, st. dev. = 2.9), the restricted and repetitive behavior subscale (mean = 4.0, st. dev. 3.3), and the ADOS total score (mean = 3.1, st. dev. = 3.1). A second dataset of neurotypical adults, the adult attention sample (n=25, 13 females, mean age = 22.8 years, st. dev. = 3.5 years). A third dataset of individuals with and without autism (n=229, 65 females; mean age = 10.45 years, st. dev. = 1.8 years; mean IQ = 113.7, st. dev. = 15.1) comprising data from ABIDE was used as an additional validation dataset. SRS scores were used from ABIDE and included the following scales: SRS total scores (mean = 42.4, st. dev. = 40.2); SRS communication (mean = 13.8, st. dev. = 14.2); SRS motivation (mean = 7.3, st. dev. = 6.9); SRS mannerisms (mean = 7.1, st. dev. = 8.5). Seventy-seven of the participants had autism. A fourth dataset of individuals (n=643, 264 females; mean age = 11.01 years, st. dev. = 2.63 years; mean IQ = 101.89, st. dev. = 16.7) comprising data from the transdiagnostic HBN sample124 was used as an additional validation dataset. In the current study, 107 had a diagnosis of autism; while 302 had a diagnosis of ADHD. Processing of these data is described elsewhere125. SRS76 total T-scores were used from HBN (mean = 56.36, st. dev. = 11.21). This sample was used to determine if the consensus model generalized to predict SRS scores. It is possible that self-selection bias may be present in the Yale youth sample and should be kept in mind. In particular, self-selection bias likely has led to the sample having a higher IQ than other studies comprising autism participants. However, the fact that the model generated from this sample to three other independently collected samples mitigates concerns about self-selection bias impacting generalizabiity.
Recruitment	Participants were recruited based on referrals from clinics at Yale and flyers posted around the community. Participants were screened over the phone for basic developmental history and MRI safety factors. Those with a history of prematurity, known genetic abnormalities, or an IQ below 70 were excluded. For the attention dataset, the relevant study was approved by the Yale University Institutional Review Board. Recruitment for the adult attention sample took place around Yale University and the surrounding community. Recruitment for ABIDE and HBN varied by site and is available for review here: https://fcon_1000.projects.nitrc.org/indi/abide/ and here: https://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network/
Ethics oversight	For the Yale youth sample and the adult attention sample, the Yale University institutional review board oversaw and approved the ethics of the study. For HBN, the study was approved by the Chesapeake Institutional Review Board. The data for ABIDE were approved by each of the 16 sites contributing data (i.e. approval was required by each of the home institutions prior to submission; see here for more details: Di Martino et al., Scientific Data, 2017).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were not pre-calculated. The sample sizes are sufficient (n = 63 x 8 scans = 504 scans in the Yale youth sample, n = 25 in the adult attention sample, n = 229 in ABIDE) given that we ensure are model generalizes in 3 separate datasets. In addition, we establish our model in the Yale youth sample (n = 504 scans), which is a fairly large sample of youth with a neurodevelopmental condition. In addition, the fact the model generalizes in larger datasets of subjects (n = 229 in ABIDE, n = 643 in HBN) suggests that we are well-powered to detect an effect in the Yale youth sample. Further, we have calculated a post-hoc power analysis: Using the R package "pwr", with a sample size of 63 subjects, and assuming a significance level of 0.05 (two-sided), a correlation between predicted and observed ADOS scores in gradCPT average of 0.445 corresponds to a power of 0.96. The use of four studies is justified, in that the first sample serves as a discovery (i.e. the Yale youth sample). Sample two (adult attention sample) is used to validate that the model from sample 1 is picking up on variance related to attention. Samples 3 and 4 are then used to validate the model in predicting autistic symptoms. Sample 2 was chosen because of the task used; sample 3 and 4

were choses because they are large datasets comprising autism (and neurotypical) participants. All four are needed to support the claims of the study.

Data exclusions

In the Yale youth sample, one hundred and two participants were scanned. We required all study participants to pass visual quality control (QC) after preprocessing (i.e. all skull-stripped data, linear registrations, and non-linear registrations were manually inspected), as well as to have data from all eight functional runs. We excluded participants with a lack of full brain coverage during functional scans (n = 7). Eight participants had an imaging artifact preventing proper registration. Fifteen participants concluded the study early. One participant was excluded due to ending the final rest run early. Participants were excluded due to a technical issue with the task laptop preventing proper recording of button presses (n = 7) and a participant (n = 1) was excluded due to pressing a combination of buttons during the gradCPT. This left 63 participants who passed visual QC and had all functional scans. In the adult attention sample, we did not exclude any of the subjects described in Rosenberg et al., from which we obtained the data (Rosenberg et al., 2016, Nature Neuroscience, https://pmc.ncbi.nlm.nih.gov/articles/PMC4696892/#S11). In the ABIDE dataset, we used the subjects described in Lake et al. (Lake et al., Biological Psychiatry 2019; .https://pmc.ncbi.nlm.nih.gov/articles/PMC7311928/). From this pool of subjects (n = 352 with low-motion data, < 0.10 mm mean frame-to-frame displacement; FFD), we additionally excluded subjects who had incomplete scan coverage (i.e., many subjects did not have adequate coverage of the cerebellum, brainstem, and dorsal portions of cortex, thus resulting in missing significant numbers of nodes from the functional parcellation used here). This resulted in the exclusion of 123 subjects with low-motion, full coverage brain data, leaving a final sample of n = 229. Of the original pool of 58 subjects with ADOS data from Lake et al.2, we excluded 25 due to a lack of full-brain coverage, resulting in a final sample size of 33 low-motion, full brain coverage subjects. Hence, we chose to analyze the SRS pool of subjects given the larger number of subjects compared to ADOS. In the HBN sample, we used the Exclusion criteria of Adkinson et al. (https://pmc.ncbi.nlm.nih.gov/articles/PMC10849571/); importantly, all subjects with functional scans with a mean FFD > 0.2 mm were excluded from the present analysis, leaving n = 643.

Replication

We used four separate datasets to ensure the model defined in dataset 1 generalized in datasets 2, 3, and 4. All attempts at replication were successful.

Randomization

Randomization was not required, as we conducted analyses on all participants in a given dataset.

Blinding

Blinding was not required, as we conducted analyses on all participants in a given dataset.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Magnetic resonance imaging

Experimental design

Design type

Both task and resting-state

Design specifications

Continuous task fMRI data were used. The gradual onset continuous performance task (gradCPT; Figure 2) has been described previously. The gradCPT tests sustained attention and inhibition and produces a range of performance scores across neurotypical and neurodiverse populations Participants viewed grayscale pictures of cities and mountains presented at the center of the screen, with images gradually transitioning from one to the next every 1,000 ms. Subjects were instructed to respond with a button press for city scenes and to withhold button presses for mountain scenes. City scenes occurred randomly 90% of the time. Performance was calculated using d' (sensitivity), the participant's hit rate minus false alarm rate. The task took 5 mins to complete; participants completed the task twice. Note that because of differences in task timing between gradCPT, the selective social attention task, and resting-state, we trimmed the gradCPT and resting-state scans were trimmed to include 4 mins of data per scan (to match the selective social task time length of 4 mins). Subjects in the adult attention sample also performed gradCPT; the same parameters were used as above, except scene transitions took 800 ms.
Participants completed a novel version of a free-viewing selective social attention (SSA) task in which an actress is presented at the center of the screen and is surrounded by objects in corners of the screen (Figure 2). Four types of clips were used where presence of speech (SP) and eye contact (EC) was manipulated. The first condition included clips in which the person smiled and made eye-contact with the camera while speaking in full sentences (e.g., "Have you ever seen a monkey? Monkeys eat bananas, swing in trees, and chase each other."; this was designated as the EC+SP+ condition). The second condition included a direct gaze condition with no speech (EC+SP-), in which the person smiled directly at the viewer while remaining silent. The third condition consisted of the person looking down at the table while

speaking in full sentences (EC-SP+). The fourth condition consisted of the person looking down at the table and not speaking (EC-SP-). Each clip lasted two minutes and was shown twice over four runs, such that eight clips were shown total. To allow successful scene transitions in between sentences, the direct gaze and speech condition lasted 2 mins 8 secs. The speech with no eye contact condition lasted 2 mins 6s. In between clips during each run, a white fixation cross on a black background was shown for 15s. Clip order was counterbalanced across participants (see Supplemental Materials for more about the counterbalancing of clips, as well as study design considerations of the Yale youth sample). Clip conditions were concatenated across runs, such that each resulting connectivity matrix comprised four minutes of data from a single scanning condition. Both gradCPT and the SSA task were presented using Psychtoolbox (version: 3.0.14; http://psychtoolbox.org/; MATLAB version R2018a) on a Lenovo IdeaPad 720S computer, with Ubuntu 16.04 LTS installed.
For resting state data in the Yale youth sample, participants viewed a fixation cross and were told to think of nothing in particular. Resting-state data were used in ABIDE; a description of the exact conditions used at each site is available here: https://fcon_1000.projects.nitrc.org/indi/abide/
In the HBN, functional images were obtained while participants completed two rest fMRI runs as well as Despicable Me and The Present movie-watching scan sessions.

Behavioral performance measures

Autism symptoms were scored in the Yale youth sample using the Autism Diagnostic Observation Schedule-2 (ADOS-2) and were ascertained by trained clinical psychologists; calibrated severity scores were used in the present work for the social affect subscale, the restricted and repetitive behavior subscale, and the ADOS total score. D-prime scores were calculated for the adult attention sample and were computed based on button presses for city scenes and withholding button presses for mountain scenes. Specifically, performance was calculated using d' (sensitivity), the participant's hit rate minus false alarm rate. In ABIDE, social responsiveness scale scores were used, which is based on parent/teacher report for those under 18 and self report for those over 19 (https://www.wpspublish.com/srs-2-social-responsiveness-scale-second-edition); the same scale was used in the HBN.

Acquisition

Imaging type(s)

Functional MRI

Field strength

3 T

Sequence & imaging parameters

For the Yale youth sample, the following imaging parameters were used: Functional images were acquired using a multiband gradient echo-planar imaging (EPI) pulse sequence with the following image parameters: 75 contiguous slices acquired in the axial-oblique plane parallel to AC-PC line, TR = 1000 ms, TE 30 ms, voxel size = 2.0 mm3, flip angle = 55 degrees, slice thickness = 2 mm, bandwidth = 1894 Hz/pixel, matrix size = 110 x 110, field of view = 220 mm, multiband factor = 5. For the adult attention sample, the following imaging parameters were used: Functional runs included 824 (task) or 363 (rest) whole-brain volumes acquired using a multiband echo-planar imaging (EPI) sequence with the following parameters: repetition time (TR) = 1000 ms, echo time (TE) = 30 ms, flip angle = 62°, acquisition matrix = 84 × 84, in-plane resolution = 2.5 mm2, 51 axial-oblique slices parallel to the ac-pc line, slice thickness = 2.5, multiband 3, acceleration factor = 2. MPRAGE parameters were as follows: TR = 2530 ms, TE = 3.32, flip angle = 7°, acquisition matrix = 256 × 256, in-plane resolution = 1.0 mm2, slice thickness = 1.0 mm, 176 sagittal slices. For the ABIDE data, imaging parameters varied by site and are available here: https://fcon_1000.projects.nitrc.org/indi/abide/. For the HBN data, imaging parameters varied by site and are available here: https://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network/

Area of acquisition

Whole brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software

https://bioimagesuiteweb.github.io/webapp/

Normalization

Linear and non-linear transformations to warp a 268-node functional atlas from Montreal Neurological Institute space to single subject space.

Normalization template

Linear and non-linear transformations to warp a 268-node functional atlas from Montreal Neurological Institute space to single subject space.

Noise and artifact removal

Functional images were motion-corrected using SPM8. Covariates of no interest were regressed from the data, including linear, quadratic, and cubic drift, a 24-parameter model of motion115, 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). Visual inspections were performed after skull-stripping, non-linear, and linear registrations to ensure there were no errors in processing.

Volume censoring

We did not censor data.

Statistical modeling & inference

Model type and settings

CPM was used to predict ADOS scores from functional connectivity data in the Yale youth sample. Briefly, using 10-fold cross-validation, connectivity matrices from a given scan condition and ADOS scores were split into an independent training set including subjects from 9 folds and a test set including the left-out fold. Linear regression was used to relate edge strength to

ADOS score in the training set. Edges most strongly associated with ADOS scores were selected (feature selection threshold of P = 0.05) for both a 'positive network' (in which increased connectivity was associated with higher ADOS scores) and a 'negative network' (in which increased connectivity was associated with lower ADOS scores). We used partial correlation to control for mean participant head motion at the feature selection step. Mean network strength was computed in both the positive and negative networks, and the difference between these network strengths was computed ('combined network strength'). A linear model was then calculated relating combined network strength to ADOS scores in the training set. In the last step, combined network strength was computed for the test set, and the model was applied to generate ADOS predictions for these unseen participants.

Effect(s) tested

Task and rest fMRI data were used to assess if any conditions resulted in predictions of ADOS measured outside of the scanner.

Specify type of analysis: ☒ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference
(See Eklund et al. 2016)

Not applicable to the analyses described here.

Correction

FDR correction was used to assess for multiple comparisons.

Models & analysis

n/a	Involved in the study
☐	☒ Functional and/or effective connectivity
☒	☐ Graph analysis
☐	☒ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Connectomes were generated using a 268-node atlas. For each subject, the mean time-course of each region of interest ("node") was computed, and the Pearson correlation coefficient was computed between each pair of nodes to achieve a symmetric 268 x 268 matrix of correlation values representing "edges" (connections between nodes) in graph theory. The Pearson correlation coefficients were then transformed to z-scores via a Fisher transformation, and only the upper triangle of the matrix was considered, yielding 35,778 unique edges.

Multivariate modeling and predictive analysis

CPM was used to predict ADOS scores from functional connectivity data in the Yale youth sample. Briefly, using 10-fold cross-validation, connectivity matrices from a given scan condition and ADOS scores were split into an independent training set including subjects from 9 folds and a test set including the left-out fold. Linear regression was used to relate edge strength to ADOS score in the training set. Edges most strongly associated with ADOS scores were selected (feature selection threshold of P = 0.05) for both a 'positive network' (in which increased connectivity was associated with higher ADOS scores) and a 'negative network' (in which increased connectivity was associated with lower ADOS scores). We used partial correlation to control for mean participant head motion at the feature selection step. Mean network strength was computed in both the positive and negative networks, and the difference between these network strengths was computed ('combined network strength'). A linear model was then calculated relating combined network strength to ADOS scores in the training set. In the last step, combined network strength was computed for the test set, and the model was applied to generate ADOS predictions for these unseen participants.