nature portfolio

Corresponding author(s): Isaac N. Treves

Last updated by author(s): Aug 27, 2025

# 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<br>*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<br>*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 | No software was used for data collection |
| Data analysis | Preprocessing and denoising was conducted in SPM12 (http://fil.ion.ucl.ac.uk/spm/) and CONN2022a (https://www.conn-toolbox.org/). Analysis was conducted in MATLAB2023a. Toolboxes included http://www.nitrc.org/projects/artifact_detect/, https://github.com/canlab/Lindquist_Dynamic_Correlation, https://github.com/swglab/CPM_CONN/, https://github.com/cocoanlab/rumination, https://github.com/canlab/CanlabCore, https://github.com/cocoanlab/cocoanCORE, and https://osf.io/m5rgu/?view_only=b987fe5766f0488db763b3f70165a096. |

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

Data will be made available upon publication at https://osf.io/m5rgu/?view_only=b987fe5766f0488db763b3f70165a096 . Code will be made available upon publication at https://osf.io/m5rgu/?view_only=b987fe5766f0488db763b3f70165a096 .

## Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

| | |
|---|---|
| Reporting on sex and gender | We recruited male and female adolescents as possible based on their self-report, with more females given the higher incidence of depressive disorders in female adolescents. (Female/Male, Study1: 51/28, 13-18 yrs.; Study2: 56/23, 12-18 yrs.; Study3: 59/57, 12-15 yrs.; Study4: 54/18, 13-18 yrs.; Study5: 64/35, 12-14 yrs.). All participants provided informed consent/ assent and were compensated hourly for participation. |
| Reporting on race, ethnicity, or other socially relevant groupings | - Study 1: 73.4% White Americans.<br>- Study 2: 62% White Americans.<br>- Study 3:82.8% White Americans.<br>- Study 4: 61% White Americans.<br>- Study 5: 81.8% White Americans.<br>See supplement for full reporting. |
| Population characteristics | - Study 1&2: Individuals scoring in elevated anhedonia range on KSADS as well as normative controls.<br>- Study 3: Individuals with parental history of MDD as well as controls.<br>- Study 4: Individuals scoring highly on rumination (CRSQ).<br>- Study 5: Individuals with parental history of MDD as well as controls.<br>See Methods for cutoffs. |
| Recruitment | Individuals were recruited from the Greater Boston Area. Participants responded to flyers posted in various neighborhoods in the city and surrounding suburbs. Participants also responded to digital flyers on Instagram and Facebook. Due to ethical considerations, advertisements targeted parents rather than children, which may have influenced the selection process of the child participants. |
| Ethics oversight | All procedures were approved by the Mass General Brigham IRB. |

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 | The sample corresponded to readily available clinical and non-clinical adolescent data with rumination scores and resting-state fMRI. At 443 adolescents, our sample size is larger than any previous published study specifically on adolescent rumination, and larger than a recent publication on adult rumination and fMRI (Kim et al., 2023). No apriori power analysis was conducted. Instead, we use predictive modelling methods with held-out data which are more robust than full-sample correlations. Research shows that multivariate fMRI markers of behaviors are replicable in datasets > 200 (Spisak et al., 2023). |
| Data exclusions | We excluded missing rumination scores as well as individuals with average framewise-displacement greater than 0.3 mm, which is a standard cutoff for adolescent fMRI research. This resulted in the exclusion of 31 adolescents for CRSQ analyses, and 23 adolescents for EMA rumination analyses. |
| Replication | We used predictive modelling with a training set and validation set consisting of Studies 1-4 split (80-20) and balanced on all covariates of interest. We also had an external dataset (Study 5). The independence of all datasets was maintained. No attempts at replication were successful. |

| | |
|---|---|
| Randomization | Not applicable |
| Blinding | Not applicable |

# 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 |
| ☒ | ☐ Plants |

## Methods

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

## Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

| | |
|---|---|
| Clinical trial registration | Secondary data analysis from trials NCT02498925 and NCT04697966. |
| Study protocol | The protocol will be made available upon publication at https://osf.io/m5rgu/?view_only=b987fe5766f0488db763b3f70165a096 . |
| Data collection | Participants completed in-person fMRI data collection procedures at the McLean Hospital Imaging Center (a.k.a. MGB Imaging at Belmont) in Belmont, Massachusetts between 2015 and 2024. All other data (e.g., self-report) were provided digitally so participants were able to complete those assessments from home or in their daily lives. |
| Outcomes | The primary outcome was the prediction performance of the adult dynamic resting-state functional connectivity model of trait rumination when applied to adolescent data, as well as linear models using DMN and whole-brain connectivity (associated with self-reported rumination in adolescents). The secondary outcomes involved an exploratory random forest analysis and testing generalizability with an independent dataset that included lower mean rumination scores. |

## Plants

| | |
|---|---|
| Seed stocks | *Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.* |
| Novel plant genotypes | *Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.* |
| Authentication | *Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.* |

## Magnetic resonance imaging

### Experimental design

| | |
|---|---|
| Design type | Resting-state |
| Design specifications | Study 1,2,3: 6.5-minutes, eyes-open<br>Study 4: three 6-minute sessions, eyes-open<br>Study 5: 5-minute session, eyes-open |

Behavioral performance measures: We administered a set of self-report questionnaires outside of the scanner, as well as ecological momentary assessments (not included in Study 5).

## Acquisition

Imaging type(s): Functional, Structural

Field strength: 3T

Sequence & imaging parameters: Images were collected on two separate scanners, one Siemens Tim Trio 3.0 Tesla MRI with a 32-channel coil (for a small subset of Study 1 & 2), and one Siemens PRISMA 3.0 Tesla MRI with a 64-channel coil (for remaining Study 1 & 2 and all of Study 3). The protocol was the same across scanners and consisted of a T1-weighted scan, one 6.5-minute resting-state scan, and a single gradient-echo fieldmap acquisition. Functional images were acquired with a multiband sequence (TR = 720 ms, TE = 30 ms, FOV = 212 mm, multiband accelerator factor = 6, voxel size = 2.5 × 2.5 × 2.5).

Study 4 Collection
Images were collected using a Siemens Prisma 3.0 Tesla MRI equipped with a 64-channel coil. The protocol consisted of a T1-weighted scan, three 6-minute, eyes-open, resting-state scans, and a single gradient-echo fieldmap acquisition. Functional images were acquired with a multiband sequence (TR = 720 ms, TE = 30 ms, FOV = 212 mm, multiband accelerator factor = 6, voxel size = 2.5 × 2.5 × 2.5).

Study 5 Collection
Images were collected using a Siemens Tim Trio 3.0 Tesla MRI equipped with a 32-channel coil. The protocol consisted of a T1-weighted scan, one five-minute eyes open resting-state scan, and a single gradient-echo fieldmap acquisition. Functional images were acquired with a multiband sequence (TR = 1300 ms, TE = 32.2 ms, FOV = 212 mm, multiband acceleration factor = 8, voxel size = 2.0 x 2.0 x 2.0).

Area of acquisition: Whole brain

Diffusion MRI ☐ Used ☒ Not used

## Preprocessing

Preprocessing software: Preprocessing was conducted in SPM12 (http://fil.ion.ucl.ac.uk/spm/) and included grey matter segmentation, realignment and unwarping with a field map, slice-timing correction, co-registration, normalization to MNI space, resampling to 2 x 2 x 2 mm voxels, and smoothing with a 4 mm FWHM gaussian filter.

Normalization: Images were normalized to MNI space using non-linear transformation

Normalization template: MNI 152

Noise and artifact removal: For denoising, we imported the SPM preprocessed data into the CONN toolbox (https://www.conn-toolbox.org/), after conducting an extra normalization step for the structural images and followed denoising steps from Kim et al. Data were band-pass filtered to include temporal frequencies between 0.008 Hz and 0.1 Hz. Nuisance covariates included image-intensity outliers from ART (i.e., "spikes"), 24 head motion parameters (six movement parameters including x, y, z, roll, pitch, and yaw, their mean-centered squares, their derivatives, and squared derivative), and the top five component scores from each of the white and cerebrospinal fluid (CSF) masks. Any participants with more than 20% of frames scrubbed, or more than 0.3 mm average framewise-displacement,47 were removed.

Volume censoring: Movement artifacts were addressed through despiking, which attenuated transient signal fluctuations in the time series prior to subject-specific first-level modeling.

## Statistical modeling & inference

Model type and settings: We used multivariate models including Lasso, CPM, and Random Forest.

Effect(s) tested: Pearson's correlation between actual and predicted scores.

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

Anatomical location(s): We used 20 DMN seeds for seed-based connectivity (Andrews-Hanna et al., 2010) and 280 whole-brain Brainnetome regions (Fan et al., 2016). We additionally used whole-brain 280 X 280 brainnetome models.

Statistic type for inference (See Eklund et al. 2016): Significance testing in training set was assessed using permutation tests shuffling participant labels 1000 times.

Correction: No correction was conducted, but corrected p-values are reported, and primary models fail to generalize even without correction.

## 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 | Dynamic conditional correlations (DCCs) are a measure of framewise connectivity between two brain regions. |
| Multivariate modeling and predictive analysis | DCCS were calculated between each seed and each Brainnetome region. The adult DMPFC-Rum model was tested in the entire dataset by using its beta weights to predict rumination scores. Then, we trained our own models using the 80-20 test split and assessing Lasso models, CPM models, and Random Forest models. If models passed our selection threshold in the training set ($p < 0.05$) , we tested them in the held-out validation set, as well as the held-out external test set. For models that generalized to held-out data (Random Forest only), we assessed their importance weights on each brainnetome region. |