nature research

Corresponding author(s): Marcel Adam Just

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   The sample size for the main analysis (17 suicidal ideators and 17 controls) matches the sizes used in our previous study (Just et.al, 2014). The high accuracy of the group membership classification (91%) indicates that the sample size was sufficient.

2. Data exclusions

   Describe any data exclusions.

   The data from 21 suicidal ideators and 24 controls were excluded from the main analysis on the basis of the pre-established minimal accuracy (0.6) of identifying which of the 30 words the participant was thinking about. The data for excluded suicidal ideators was included in the further analyses.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   The results of the main analysis were replicated for the initially excluded suicidal ideators.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   The suicidal ideator participants were recruited from a clinically verified population with current suicidal ideation. The control group was recruited from the community with the stipulation that there was no personal or family history of psychiatric disorder or suicide attempt.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   The investigators were not blinded to the group membership of participants; however, the machine learning classifier was not given the group membership information for the test participant.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

Matlab version R2014a
SPM8
Custom Matlab code (most of this code was used in the previously published studies)

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eucaryotic cell lines were used.

b. Describe the method of cell line authentication used.

No eucaryotic cell lines were used.

c. Report whether the cell lines were tested for mycoplasma contamination.

No eucaryotic cell lines were used.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No commonly misidentified cell lines were used.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	The covariates in the two groups are reported in detail and treated statistically in the manuscript.

nature research

Corresponding author(s): Marcel Adam Just

☐ Initial submission ☐ Revised version ☒ Final submission

MRI Studies Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Experimental design

1. Describe the experimental design.	event-related design
2. Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.	30 concepts were presented in 6 blocks, each block in a different random order. Each concept was displayed for 3 sec followed by a 4 sec blank interval. 17 sec long fixation intervals were included periodically to provide an activation baseline
3. Describe how behavioral performance was measured.	The behavioral performance was not measured during the experiment; the performance was estimated by the sufficiently high accuracy of the machine learning classifier that identified which of the 30 concepts the participant was thinking about.

Acquisition

4. Imaging	
a. Specify the type(s) of imaging.	Functional MRI
b. Specify the field strength (in Tesla).	3T
c. Provide the essential sequence imaging parameters.	Echo-planar pulse sequence, TR (repetition time) = 1000 ms, TE (echo time) = 30 ms, flip angle = 60 degrees, FOV (field of view) = 20 cm, matrix size = 64 x 64, voxel size of 3.125 x 3.125 x 5 mm thick (skipping 1mm between slices), 20 AC-PC aligned brain slices
d. For diffusion MRI, provide full details of imaging parameters.	No diffusion MRI was acquired.
5. State area of acquisition.	The acquisition matrix covered all of the cerebrum.

Preprocessing

6. Describe the software used for preprocessing.	fMRI images were preprocessed with SPM8 (Wellcome Dept. of Cog. Neurology). The images were slice-timing- and motion-corrected, and spatially normalized to the MNI template without changing voxel size (3.125 x 3.125 x 6 mm)
7. Normalization	
a. If data were normalized/standardized, describe the approach(es).	Non-linear normalization to the functional MNI template
b. Describe the template used for normalization/ transformation.	MNI305
8. Describe your procedure for artifact and structured noise removal.	Motion parameters were estimated and corrected in original subject's space
9. Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.	The main analysis was performed on activation in the clusters of stable voxels. These clusters were identified algorithmically based exclusively on the training data inside the cross-validation folds.

Statistical modeling & inference

10. Define your model type and settings.	Multivariate analysis using machine learning classifier (Gaussian Naïve Bayes)
11. Specify the precise effect tested.	The main effect tested was the ability to identify group membership of a participant (suicidal ideator or control), based on the fMRI activation during processing of discriminative concepts.
12. Analysis	
a. Specify whether analysis is whole brain or ROI-based.	Whole-brain analysis restricted to the clusters of stable voxels that were identified algorithmically using the training data.
b. If ROI-based, describe how anatomical locations were determined.	No anatomical ROIs were used.
13. State the statistic type for inference. (See Eklund et al. 2016.)	The classifier performance was compared to the accuracy expected by chance (obtained using binomial distribution or random classifier).
14. Describe the type of correction and how it is obtained for multiple comparisons.	The main analysis uses cross-validated machine learning classification that does not require statistical correction.
15. Connectivity	
a. For functional and/or effective connectivity, report the measures of dependence used and the model details.	No connectivity analyses were performed.
b. For graph analysis, report the dependent variable and functional connectivity measure.	No graph analyses were performed.
16. For multivariate modeling and predictive analysis, specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.	In the main analysis, the independent variables (features) were the activation levels for a number of (discriminating) concepts in a set of (discriminating) brain locations. The classifier was trained on the data of all but one participant, and the group membership of the left out participant was predicted. The main evaluation metric was the accuracy of the classification.