Wu and Zhao (2021)

EPPI-Centre (2003) & Critical Appraisal Skills Programme (2018)

If the study has a broad focus and this data extraction focuses on just one component of the study, please specify this here

☒ Not applicable (whole study is focus of data extraction)

☐ Specific focus of this data extraction (please specify)

Study aim(s) and rationale

Was the study informed by, or linked to, an existing body of empirical and/or theoretical research?

Please write in authors' declaration if there is one. Elaborate if necessary, but indicate which aspects are reviewers' interpretation.

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- Stereotype threat + the fact that at that time only two studies using an fMRI to uncover the neural mechanisms of stereotype threat were available.
- RSDC : RS-fMRI degree centrality is a graph theory-based network analysis of the number of edges connecting to a node or the node strength for a given node (a voxel) in a whole-brain network.

Do authors report how the study was funded?

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

This study was supported by the National Natural Science Foundation of China (31371055) and the Major Project for Key Research Institutes of Humanities and Social Science by the Ministry of Education (16JJD1900007).

Study research question(s) and its policy or practice focus

What is/are the topic focus/foci of the study?

The current study aimed to investigate the neffects of ST on RSDC under math ST.

What is/are the population focus/foci of the study?

- women

What is the relevant age group?

☐ Not applicate (focus not learners)

☐ 0 - 4

☐ 5 - 10

☐ 11 - 16

☐ 17 - 20

☐ 21 and over

☒ Not stated/unclear

What is the sex of the population focus/foci?

☐ Not applicate (focus not learners)

☒ Female only

☐ Male only

☐ Mixed sex

☐ Not stated/unclear

What is/are the educational setting(s) of the study?

☐ Community centre

☐ Correctional institution

☐ Government department

☒ Higher education institution

☐ Home

☐ Independent school

☐ Local education authority

☐ Nursery school

☐ Other early years setting

☐ Post-compulsory education institution

☐ Primary school

☐ Residential school

☐ Secondary school

☐ Special needs school

☐ Workplace

☐ Other educational setting

In Which country or cuntries was the study carried out?

☒ Explicitly stated (please specify)

☐ Not stated/unclear (please specify)

- China

Please describe in more detail the specific phenomena, factors, services, or interventions with which the study is concerned

What are the study reserach questions and/or hypotheses?

Research questions or hypotheses operationalise the aims of the study. Please write in authors' description if there is one. Elaborate if necessary, but indicate which aspects are reviewers' interpretation.

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- Effects of ST on RSDC under math ST.
- We speculated that the RSDC of brain regions related to the regulation of social emotions (e.g. the medial prefrontal cortex and anterior cingulate cortex) might be increased.
- Considering the roles of the hippocampus in the generation of stressful responses, we hypothesized that the RSDC of the hippocampus might be decreased under ST, which might make the individuals more prone to experience stress-based arousal.
- Except in this region, due to the self-relevant processes being increased by ST, we thought that the RSDC of the brain regions related to self-memory should be increased. Therefore, we further speculated that the RSDC of the brain regions, such as the right posterior parietal regions (PPC), related to the retrieval of episodic memory or autobiographical memory might be increased.

Methods - Design

Which variables or concepts, if any, does the study aim to measure or examine?

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- Stereotype threat manipulation
- fMRI
- RS-fMRI degree centrality (RSDC)
- Math problems

Study timing

Please indicate all that apply and give further details where possible.

If the study examines one or more samples, but each at only one point in time it is cross-sectional.
If the study examines the same samples, but as they have changed over time, it is retrospective, provided that the interest is in starting at one timepoint and looking backwards over time.
If the study examines the same samples as they have changed over time and if data are collected forward over time, it is prospective provided that the interest is in starting at one timepoint and looking forward in time.

☒ Cross-sectional

☐ Retrospective

☐ Prospective

☐ Not stated/unclear (please specify)

If the study is an evaluation, when were measurements of the variable(s) used for outcome made, in relation to the intervention?

If at least one of the outcome variables is measured both before and after the intervention, please use the before and after category.

☐ Not applicable (not an evaluation)

☒ Before and after

☐ Only after

☐ Other (please specify)

☐ Not stated/unclear (please specify)

Methods - Groups

If comparisons are being made between two or more groups, please specify the basis of any divisions made for making these comparisons.

Please give further details where possible.

☐ Not applicable (not more than one group)

☒ Prospecitive allocation into more than one group (e.g. allocation to different interventions, or allocation to intervention and control groups)

☐ No prospective allocation but use of pre-existing differences to create comparison groups (e.g. receiving different interventions, or characterised by different levels of a variable such as social class)

☐ Other (please specify)

☐ Not stated/unclear (please specify)

How do the groups differ?

☐ Not applicable (not more than one group)

☒ Explicityly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- control vs st manipulation

Number of groups

For instance, in studies in which comparisons are made between groups, this may be the number of groups into which the dataset is divided for analysis (e.g. social class, or form size), or the number of groups allocated to, or receiving, an intervention.

☐ Not applicable (not more than one group)

☐ One

☒ Two

☐ Three

☐ Four or more (please specify)

☐ Other/unclear (please specify)

Was the assignment of participants to interventions randomised?

☐ Not applicable (not more than one group)

☐ Not applicate (no prospective allocation)

☒ Random

☐ Quasi-random

☐ Non-random

☐ Not stated/unclear (please specify)

Where there was prospective allocation to more than one group, was the allocation sequence concealed from participants and those enrolling them until after enrolment?

Bias can be introduced, consciously or otherwise, if the allocation of pupils or classes or schools to a programme or intervention is made in the knowledge of key characteristics of those allocated. For example: children with more serious reading difficulty might be seen

as in greater need and might be more likely to be allocated to the 'new' programme, or the opposite might happen. Either would introduce bias.

☐ Not applicable (not more than one group)

☐ Not applicable (no prospective allocation)

☒ Yes (please specify)

☐ No (please specify)

☐ Not stated/unclear (please specify)

- ST group was introduced to the material with statements such as "women are bad at math across all cultures"

- control group was instructed to read a scientific investigation about two fictitious mountain peaks

Apart from the experimental intervention, did each study group receive the same level of care (that is, were they treated equally)?

☒ Yes
☐ No
☐ Can't tell

Study design summary

In addition to answering the questions in this section, describe the study design in your own words. You may want to draw upon and elaborate the answers you have already given.

Methods - Sampling strategy

Are the authors trying to produce findings that are representative of a given population?

Please write in authors' description. If authors do not specify please indicate reviewers' interpretation.

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- women under math stereotype threat

Which methods does the study use to identify people or groups of people to sample from and what is the sampling frame?

e.g. telephone directory, electoral register, postcode, school listing, etc. There may be two stages – e.g. first sampling schools and then classes or pupils within them.

☐ Not applicable (please specify)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)
☒ Not stated/unclear (please specify)

Which methods does the study use to select people or groups of people (from the sampling frame)?

e.g. selecting people at random, systematically - selecting for example every 5th person, purposively in order to reach a quota for a given characteristic.

☐ Not applicable (no sampling frame)

☐ Explicitly stated (please specify)

☒ Implicit (please specify)

☐ Not stated/unclear (please specify)

- females from the Southwest University of China

Planned sample size

If more than one group please give details for each group separately.

☐ Not applicable (please specify)
☐ Explicitly stated (please specify)
☒ stated/unclear (please specify)

Methods - Recruitment and consent

Which methods are used to recruit people into the study?

e.g. letters of invitation, telephone contact, face-to-face contact.

☐ Not applicable (please specify)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)
- [x Not stated/unclear (please specify)

Were any incentives provided to recruit people into the study?

☐ Not applicable (please specify)
☐ Explicitly stated (please specify)
☒ Not stated/unclear (please specify)

Was consent sought?

Please comment on the quality of consent if relevant.

☐ Not applicable (please specify)
☒ Participant consent sought
☐ Parental consent sought
☐ Other consent sought
☐ Consent not sought
☐ Not stated/unclear (please specify)

Are there any other details relevant to recruitment and consent?

☐ No
☒ Yes (please specify)

All the subjects gave written informed consent. The studies involving human participants were reviewed and approved by Brain Imaging Center Institutional Review Board of Southwest China University. The patients/participants provided their written informed consent to participate in this study.

Methods - Actual sample

What was the total number of participants in the study (the actual sample)?

If more than one group is being compared please give numbers for each group.

☐ Not applicable (e.g. study of policies, documents, etc)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)
☒ Not stated/unclear (please specify)

What is the proportion of those selected for the study who actually participated in the study?

Please specify numbers and percentages if possible.

☐ Not applicable (e.g. study of policies, documents, etc)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)
☒ Not stated/unclear (please specify)

Which country/countries are the individuals in the actual sample from?

If UK, please distinguish between England, Scotland, N. Ireland, and Wales if possible. If from different countries, please give numbers for each. If more than one group is being compared, please describe for each group.

☐ Not applicable (e.g. study of policies, documents, etc)

☐ Explicitly stated (please specify)

☒ Implicit (please specify)

☐ Not stated/unclear (please specify)

- Study takes place at a Chinese University, no further details given.

What ages are covered by the actual sample?

Please give the numbers of the sample that fall within each of the given categories. If necessary, refer to a page number in the report (e.g. for a useful table). If more than one group is being compared, please describe for each group. If follow-up study, age at entry to the study.

☐ Not applicable (e.g. study of policies, documents, etc)

☐ 0 to 4

☐ 5 to 10

☐ 11 to 16

☒ 17 to 20

☒ 21 and over

☐ Not stated/unclear (please specify)

- mean age was 20.75, SD = 1.79 years

What is the socio-economic status of the individuals within the actual sample?

If more than one group is being compared, please describe for each group.

☐ Not applicable (e.g. study of policies, documents, etc)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)
☒ Not stated/unclear (please specify)

What is the ethnicity of the individuals within the actual sample?

If more than one group is being compared, please describe for each group.

☐ Not applicable (e.g. study of policies, documents, etc)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)
☒ Not stated/unclear (please specify)

What is known about the special educational needs of individuals within the actual sample?

e.g. specific learning, physical, emotional, behavioural, intellectual difficulties.

☐ Not applicable (e.g. study of policies, documents, etc)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)

☒ Not stated/unclear (please specify)

Is there any other useful information about the study participants?

☐ Not applicable (e.g. study of policies, documents, etc)
☐ Explicitly stated (please specify no/s.)
☐ Implicit (please specify)
☒ Not stated/unclear (please specify)

How representative was the achieved sample (as recruited at the start of the study) in relation to the aims of the sampling frame?

Please specify basis for your decision.

☐ Not applicable (e.g. study of policies, documents, etc)
☐ Not applicable (no sampling frame)
☐ High (please specify)
☒ Medium (please specify)
☐ Low (please specify)
☐ Unclear (please specify)

If the study involves studying samples prospectively over time, what proportion of the sample dropped out over the course of the study?

If the study involves more than one group, please give drop-out rates for each group separately. If necessary, refer to a page number in the report (e.g. for a useful table).

☐ Not applicable (e.g. study of policies, documents, etc)
☒ Not applicable (not following samples prospectively over time)
☐ Explicitly stated (please specify)
☐ Implicit (please specify)
☐ Not stated/unclear

For studies that involve following samples prospectively over time, do the authors provide any information on whether and/or how those who dropped out of the study differ from those who remained in the study?

☐ Not applicable (e.g. study of policies, documents, etc)
☒ Not applicable (not following samples prospectively over time)
☐ Not applicable (no drop outs)
☐ Yes (please specify)
☐ No

If the study involves following samples prospectively over time, do authors provide baseline values of key variables such as those being used as outcomes and relevant socio-demographic variables?

☐ Not applicable (e.g. study of policies, documents, etc)
☒ Not applicable (not following samples prospectively over time)
☐ Yes (please specify)

☐ No

Methods - Data collection

Please describe the main types of data collected and specify if they were used (a) to define the sample; (b) to measure aspects of the sample as findings of the study?

☐ Details

- math problems, fMRI, RS-fMRI degree centrality (RSDC) -> b
- manipulation check

Which methods were used to collect the data?

Please indicate all that apply and give further detail where possible.

☐ Curriculum-based assessment
☐ Focus group
☐ Group interview
☐ One to one interview (face to face or by phone)
☐ Observation
☐ Self-completion questionnaire
☐ Self-completion report or diary
☒ Exams
☒ Clinical test
☐ Practical test
☒ Psychological test
☐ Hypothetical scenario including vignettes
☐ School/college records (e.g. attendance records etc)
☐ Secondary data such as publicly available statistics
☐ Other documentation
☐ Not stated/unclear (please specify)

Details of data collection methods or tool(s).

Please provide details including names for all tools used to collect data and examples of any questions/items given. Also please state whether source is cited in the report.

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- ST manipulation: e.g. “women are abd at math across all cultures”
- Control group: scientific investigation about two fictitious mountain peaks
- Math problems: There were three numbers in each math problem. The subjects were asked to calculate whether the result of the subtraction of the first and second numbers

could be divided by the third number, e.g.: whether the result of "(30 - 26) / 7 = ?" was an integer.

- Manipulation check: 7-point scale (from strongly disagree to strongly agree), "I am worried that the experimenters will conclude that women are bad at math based on my performance."

- ST and Control material was verified with a different group of 112 female subjects of the same university (56 for each ST/control group). They read the ST or control materials, then completed 20 math problems.

- RS-fMRI scan, each included 242 scans with 484 s in duration.

- 20 math problems

- whole-brain RS-fMRI images were acquired from a Siemens 3T scanner (MAGENTOM Trio, a Tim system) with a gradient-echo echo-planar imaging sequence. For each subjest, 242 functional images were acquired. Additionally, high-resolution T1-weighted anatomical images were acquired using a magnetization-prepared, rapid gradient echo sequence.

- RS-fMRI images were pre-processes using statistical parametric mapping software (SPM8), a toolbox for Data Processing and Analysis for Brain Imaging (DPABI). DICOM data were converted to NIFTI images, first 10 images were discarded. All functional images were normalized to the Montreal Neurological Institute (MNI) space of 3 x 3 x 3 mm voxel sized using the segmented data. Normalized images were spatially smoothed with a Gaussian kernel having a full width at half maximum (FWHM) of 8 mm.

- RSDC was calculated using DPABI. RSDC values of each region of interest (ROI) were extracted by using the "ROI Signal Extractor" of DPABI.

Who collected the data?

Please indicate all that apply and give further detail where possible.

☒ Researcher
☐ Head teacher/Senior management
☐ Teaching or other staff
☐ Parents
☐ Pupils/students
☐ Governors
☐ LEA/Government officials
☐ Other education practitioner
☐ Other (please specify)
☐ Not stated/unclear

Do the authors describe any ways they addressed the reliability of their data collection tools/methods?

e.g. test-retest methods (Where more than one tool was employed please provide details for each.)

☐ Details

Do the authors describe any ways they have addressed the validity of their data collection tools/methods?

e.g. mention previous validation of tools, published version of tools, involvement of target population in development of tools. (Where more than one tool was employed please provide details for each.)

☒ Details

ST and Control Material Verification was performed with a sample of 112 female subjects selected from the Southwest Univeristy in China.

Was there concealment of study allocation or other key factors from those carrying out measurement of outcome – if relevant?

Not applicable – e.g. analysis of existing data, qualitative study. No – e.g. assessment of reading progress for dyslexic pupils done by teacher who provided intervention. Yes – e.g. researcher assessing pupil knowledge of drugs - unaware of pupil allocation.

☐ Not applicable (please say why)
☐ Yes (please specify)
☒ No (please specify)

Not stated

Where were the data collected?

e.g. school, home.

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Unclear/not stated (please specify)

- MRI lab

Are there other important features of data collection?

e.g. use of video or audio tape; ethical issues such as confidentiality etc.

☐ Details

Methods - Data analysis

Which methods were used to analyse the data?

Please give details e.g. for in-depth interviews, how were the data handled? Details of statistical analysis can be given next.

☒ Explicitly stated (please specify)

☐ Implicit (please specify)

☐ Not stated/unclear (please specify)

- see above

Which statistical methods, if any, were used in the analysis?

☐ Details

- Mixed-effect analysis: 2 (test: pre-test vs. post-test) x 2 (ST: ST group vs. control group) mixed-effect analysis for binary and weighted graphs
- Statistical criterion was set at a Gaussian random field (GRF) corrected threshold for the main effect of the test, the main effect of ST, and the interaction effect between ST and test.

- Interaction between ST and MC score predicting RSDC: the interactions between ST and MC scores were assessed by using t-contrast, where the statistical inference was $p < 0.0005$ at the vox level with GRF-corrected $p < 0.05$ at the cluster level.
- Results of the MC: two-sample t-test

What rationale do the authors give for the methods of analysis for the study?

e.g. for their methods of sampling, data collection, or analysis.

☐ Details

For evaluation studies that use prospective allocation, please specify the basis on which data analysis was carried out.

'Intention to intervene' means that data were analysed on the basis of the original number of participants as recruited into the different groups. 'Intervention received' means data were analysed on the basis of the number of participants actually receiving the intervention.

☐ Not applicable (not an evaluation study with prospective allocation)
☐ 'Intention to intervene'
☒ 'Intervention received'
☐ Not stated/unclear (please specify)

Do the authors describe any ways they have addressed the reliability of data analysis?

e.g. using more than one researcher to analyse data, looking for negative cases.

☐ Details

Do the authors describe any ways they have addressed the validity of data analysis?

e.g. internal or external consistency; checking results with participants.

☐ Details

Do the authors describe strategies used in the analysis to control for bias from confounding variables?

☐ Details

Please describe any other important features of the analysis.

☐ Details

Please comment on any other analytic or statistical issues if relevant.

☐ Details

Results and Conclusions

How are the results of the study presented?

e.g. as quotations/figures within text, in tables, appendices.

☐ Details

- figures, tables, in text

What are the results of the study as reported by authors?

Please give details and refer to page numbers in the report(s) of the study where necessary (e.g. for key tables).

☐ Details

Results of the Manipulation Check (MC):
- MC score for the ST group was 4.52 (SD: 1.58), while the MC score for the control group was 3.47 (SD: 1.31). - Two-sample t-test analysis showed that the mean MC score for the ST group was significantly higher than that of the control group.

Results of RSDC Analysis:
Mixed-Effect Analysis of the Binary Graph:
- Results of 2 (test) x 2 (ST) mixed-effect analysis for the binary graph shpwed that the main effect of the test was significant to the left hippocampus, middle cingulate gyrus (MCG), right cerebellum, and left precentral gyrus (PCG), while the interaction between the test and ST was significant in the left cerebellum anterior lobe, left hippocampus, left precuneus, and left MOC (see Figure 1, Table 2). - Brain regions that showed a significant main effect of the test, the results showed that the mean RSDC z-value only in MCG was higher for the ST groups relative to the control group. - Of those brain regions that had significant

interactions, the mean RSDC z-value in the left cerebellum was lower for the ST group relative to the control group; the mean RSDC z-value in the right superior parietal gyrus (SPG) was higher for the ST group relative to the control group; the mean RSDC z-value in the left precuneus was higher for the ST group relative to the control group; the mean RSDC z-value in the left MOG was higher for the ST group relative to the control group; the mean RSDC z-value in the right AG was higher for the ST group relative to the control group; and the mean RSDC z-value in the left hippocampus was lower for the ST group relative to the control group. - The RSDC for these regions was not significantly correlated with the MC score (see Table 2).

Mixed-effect analysis of the weighted graph:
- The results of whole brain 2 (test) x 2 (ST) mixed-effect analysis for the weighted graph showed that the main effect of the test was significant in the left hippocampus, left MCG, right cerebellum, and left PCFG, while the interaction between the test and ST was significant in the left cerebellum anterior lobe, left precuneus, left MCG, right SPG, and right AG (see Figure 2, Table 2) - For the brain regions that had a significant main effect of the test, the results showed that the mean RSDC z-value only in MCG was higher for the ST group relative to the control group. - For the brain regions that had significant interactions, the mean RSDC z-value in the left cerebellum was lower for the ST group relative to the control group; the mean RSDC z-value in the left precuneus was higher for the ST group relative to the control group; the mean RSDC z-value in the right SPG was higher for the ST group relative to the control group; the mean RSDC z-value wasin the left precuneus was higher for the ST group relative to the control group; the mean RSDC z-value in the left MOG was higher for the ST group relative to the control group; and the mean RSDC z-value in AG was higher for the ST group relative to the control group. - However, the RSDC for these regions was not significantly correlated with the MC score (see table 3).

The modulating roles of ST in the relationships between RSDC and MC score:
- The interaction between ST and MC score was significant in the right anterior temporal lobe (ATL) and the right hippocampus/amygdala for the binary graph, while only the right hippocampus/amygdala showed significant results for the weighted graph (see figure 3) - for the binary graph, the results showed that the interaction between ST and psot-test RSDC fo the right ATL was significant - the sample slope test showed that the post-test RSDC for the ST group was negatively correlated with the MC score, while the post-test RSDC for the ST group was positively correlated with the MC score. - The interaction between ST and post-test RSDC of the right hippocampus was significant - The sample slope test showed that the post-test RSDC of the right hippocampus was negatively correlated with the MC score in the ST group, while the post-test RSC of the hippocampus was positively correlated with the MC score in the control group, see Figure 4 - For the weighted graph, the results showed that the interaction between ST and post-test RSDC of the right ATL was significant for the weighted graph. - the sample slope test showed that the post-test RSDC for the ST group was negatively correlated with the MC score, while the post-test RSDC for the ST group was positively correlated with the MC score. - The interaction between ST and post-test RSDC of the right hippocampus was significant - The sample slope test showed that the post-test RSDC of the right hippocampus was negatively correlated with the MC score in the ST group, while the post-test RSDC of the hippocampus was positively

correlated with the MC score in the control group, see Figure 4

Was the precision of the estimate of the intervention or treatment effect reported?

- CONSIDER:
  - Were confidence intervals (CIs) reported?

☐ Yes
☒ No
☐ Can't tell

Are there any obvious shortcomings in the reporting of the data?

☐ Yes (please specify)
☒ No

Do the authors report on all variables they aimed to study as specified in their aims/research questions?

This excludes variables just used to describe the sample.

☒ Yes (please specify)
☐ No

Do the authors state where the full original data are stored?

☒ Yes (please specify)

☐ No

- The datasets generated for this study are available on request to the corresponding author.

What do the author(s) conclude about the findings of the study?

Please give details and refer to page numbers in the report of the study where necessary.

☐ Details

Results showed that the RSDC decreased in the left MOG and left hippocampus, while the RSDC increased in the right MCG, right SPG, right AG, and left precuneus. Furthermore, the results also showed that the right ATL and right hippocampus were negatively correlated with the MC score in the ST group, while the right ATL and the right hippocampus were positively correlated with the MC score in the control group.

We speculated that the increased RSDC in the left precuneus and MCG might reflect that some self-relevant thoughts are generated by retrieving autobiographical memory or episodic memory.
The self-relevant processes under ST could also be reflected by the increased RSDC in the PPC.
According to the results, the RSDC of the left hippocampus was decreased, and the RSDC of the right hippocampus was negatively correlated with the MC score under the ST condition.

These results might reflect that the importance of the hippocampus in the brain networks was decreased by the ST condition, which might make individuals prone to exhibit stress-based arousal.
According to the result, the RSDC of the right ATL was negatively correlated with the MC score in the ST group, while the RSDC of the rigth ATL was positively correlated with the MC score in the control group. These findings reflect that the relative importance of the brain regions related to social concepts is decreased under ST, especially for more threatened individuals. It is dangerous to increase the self-relevant processes related to ST and at the same time decrease the functions of social concepts. Doing so might make the indiviual more prone to be influenced by their stigmatized social identity and discount other beneficial social identities.

Conclusion:
We detected the effects of ST on brain network degree centrality by directly comparing the RSDC between ST and control groups. These findings expand the knowledge of the neural basis of ST. Specifically, the results suggest that ST decreases the imporance of brain regions related to social concepts and stress regulation in the whole brain networks and, at the same time, increases the importance of brain regions associated with self-relevant processes.

Limitations:
We acknowledge several limitations of the study. First, the performance of the math problem was not measured in an RS-fMRI study, and we did not confirm whether the altered RSDCs were related to underperformance under ST. Second, Steele and Aronson (1995) found that ST could lead to some self-concerning processes, such as self-doubts, self-validating the stereotype, and stereotype avoidance. Although we speculate that the altered RSDCs might be ralted to these mental processes, we could not provide direct evidence from them because behaviour related to these processes was not measured. Third, due to this evidence being exclusively from female math ST, future studies should verify whether these results generalize to other STs (e.g. racial ST or ST related to social status). Therefore, we suggest that a sophisticated experiment integrating behavioural, task, and RS-fMRI may be required to reveal the detailed mechanism of ST.

Quality of the study - Reporting

Is the context of the study adequately described?

Consider your answer to questions: Why was this study done at this point in time, in those contexts and with those people or institutions? (Section B question 2) Was the study informed by or linked to an existing body of empirical and/or theoretical research? (Section B question 3) Which of the following groups were consulted in working out the aims to be addressed in the study? (Section B question 4) Do the authors report how the study was funded? (Section B question 5) When was the study carried out? (Section B question 6)

☒ Yes (please specify)
☐ No (please specify)

Are the aims of the study clearly reported?

Consider your answer to questions: What are the broad aims of the study? (Section B question 1) What are the study research questions and/or hypotheses? (Section C question 10)

☒ Yes (please specify)
☐ No (please specify)

Is there an adequate description of the sample used in the study and how the sample was identified and recruited?

Consider your answer to all questions in Methods on 'Sampling Strategy', 'Recruitment and Consent', and 'Actual Sample'.

☒ Yes (please specify)
☐ No (please specify)

Is there an adequate description of the methods used in the study to collect data?

Consider your answer to the following questions in Section I: Which methods were used to collect the data? Details of data collection methods or tools Who collected the data? Do the authors describe the setting where the data were collected? Are there other important features of the data collection procedures?

☒ Yes (please specify)
☐ No (please specify)

Is there an adequate description of the methods of data analysis?

Consider your answer to the following questions in Section J: Which methods were used to analyse the data? What statistical methods, if any, were used in the analysis? Who carried out the data analysis?

☒ Yes (please specify)
☐ No (please specify)

Is the study replicable from this report?

☒ Yes (please specify)
☐ No (please specify)

Do the authors avoid selective reporting bias?

(e.g. do they report on all variables they aimed to study as specified in their aims/research questions?)

☒ Yes (please specify)
☐ No (please specify)

Quality of the study - Methods and data

Are there ethical concerns about the way the study was done?

Consider consent, funding, privacy, etc.

☐ Yes, some concerns (please specify)
☒ No concerns

Were students and/or parents appropriately involved in the design or conduct of the study?

☒ Yes, a lot (please specify)
☐ Yes, a little (please specify)
☐ No (please specify)

Is there sufficient justification for why the study was done the way it was?

☒ Yes (please specify)
☐ No (please specify)

Was the choice of research design appropriate for addressing the research question(s) posed?

☒ Yes (please specify)
☐ No (please specify)

To what extent are the research design and methods employed able to rule out any other sources of error/bias which would lead to alternative explanations for the findings of the study?

e.g. (1) In an evaluation, was the process by which participants were allocated to or otherwise received the factor being evaluated concealed and not predictable in advance? If not, were sufficient substitute procedures employed with adequate rigour to rule out any alternative explanations of the findings which arise as a result? e.g. (2) Was the attrition rate low and if applicable similar between different groups?

☒ A lot (please specify)
☐ A little (please specify)
☐ Not at all (please specify)

How generalisable are the study results?

☐ Details

- just towards female math ST

Weight of evidence - A: Taking account of all quality assessment issues, can the study findings be trusted in answering the study question(s)?

In some studies it is difficult to distinguish between the findings of the study and the conclusions. In those cases please code the trustworthiness of this combined results/conclusion.

Please remember to complete the weight of evidence questions B-D which are in your review specific data extraction guidelines.

☒ High trustworthiness (please specify)
☐ Medium trustworthiness (please specify)
☐ Low trustworthiness (please specify)

Have sufficient attempts been made to justify the conclusions drawn from the findings so that the conclusions are trustworthy?

☐ Not applicable (results and conclusions inseparable)
☒ High trustworthiness
☐ Medium trustworthiness
☐ Low trustworthiness

Wells et al. (2014)

CASE CONTROL STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

Is the case definition adequate?

- a) yes, with independent validation
- b) yes, e.g., record linkage or based on self reports
- c) no description

Representativeness of the cases

- a) consecutive or obviously representative series of cases *
- b) potential for selection biases or not stated

Selection of Controls

- a) community controls *
- b) hospital controls
- c) no description

Definition of Controls

- a) no history of disease (endpoint) *
- b) no description of source

Comparability

Comparability of cases and controls on the basis of the design or analysis

- a) study controls for _____ (Select the most important factor.) *
- b) study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.)

Exposure

Ascertainment of exposure

- a) secure record (e.g., surgical records) *
- b) structured interview where blind to case/control status *
- c) interview not blinded to case/control status
- d) written self report or medical record only
- e) no description

Same method of ascertainment for cases and controls

- a) yes *
- b) no

Non-Response rate

- a) same rate for both groups *
- b) non respondents described
- c) rate different and no designation

COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability.

Selection

Representativeness of the exposed cohort

- a) truly representative of the average _____ (describe) in the community *
- b) somewhat representative of the average _____ in the community *
- c) selected group of users, e.g., nurses, volunteers
- d) no description of the derivation of the cohort

Selection of the non exposed cohort

- a) drawn from the same community as the exposed cohort *
- b) drawn from a different source
- c) no description of the derivation of the non exposed cohort

Ascertainment of exposure

- a) secure record (e.g., surgical records) *
- b) structured interview *
- c) written self report
- d) no description

Demonstration that outcome of interest was not present at start of study

- a) yes *
- b) no

Comparability

Comparability of cohorts on the basis of the design or analysis

- a) study controls for _____ (select the most important factor) *
- b) study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.)

Outcome

Assessment of outcome

- a) independent blind assessment *
- b) record linkage *
- c) self report
- d) no description

Was follow-up long enough for outcomes to occur

- a) yes (select an adequate follow up period for outcome of interest) *
- b) no

Adequacy of follow up of cohorts

- a) complete follow up - all subjects accounted for *
- b) subjects lost to follow up unlikely to introduce bias - small number lost - > _____ % (select an adequate %) follow up, or description provided of those lost) *
- c) follow up rate < _____% (select an adequate %) and no description of those lost
- d) no statement

University of Glasgow (n.d.)

DOES THIS REVIEW ADDRESS A CLEAR QUESTION?

Did the review address a clearly focussed issue?

- Was there enough information on:
  - The population studied
  - The intervention given
  - The outcomes considered

☐ Yes
☐ Can't tell
☐ No

Did the authors look for the appropriate sort of papers?

- The 'best sort of studies' would:
  - Address the review's question
  - Have an appropriate study design

☐ Yes
☐ Can't tell
☐ No

ARE THE RESULTS OF THIS REVIEW VALID?

Do you think the important, relevant studies were included?

- Look for:
  - Which bibliographic databases were used
  - Follow up from reference lists
  - Personal contact with experts
  - Search for unpublished as well as published studies
  - Search for non-English language studies

☐ Yes
☐ Can't tell
☐ No

Did the review's authors do enough to assess the quality of the included studies?

- The authors need to consider the rigour of the studies they have identified. Lack of rigour may affect the studies results.

☐ Yes
☐ Can't tell
☐ No

If the results of the review have been combined, was it reasonable to do so?

- Consider whether:
  - The results were similar from study to study
  - The results of all the included studies are clearly displayed

  - The results of the different studies are similar
  - The reasons for any variations are discussed

☐ Yes
☐ Can't tell
☐ No

WHAT ARE THE RESULTS?

What is the overall result of the review?

- Consider:
  - If you are clear about the review's 'bottom line' results
  - What these are (numerically if appropriate)
  - How were the results expressed (NNT, odds ratio, etc)

How precise are the results?

- Are the results presented with confidence intervals?

☐ Yes
☐ Can't tell
☐ No

WILL THE RESULTS HELP LOCALLY?

Can the results be applied to the local population?

- Consider whether:
  - The patients covered by the review could be sufficiently different from your population to cause concern
  - Your local setting is likely to differ much from that of the review

☐ Yes
☐ Can't tell
☐ No

Were all important outcomes considered?

☐ Yes
☐ Can't tell
☐ No

Are the benefits worth the harms and costs?

- Even if this is not addressed by the review, what do you think?

☐ Yes
☐ Can't tell
☐ No

References

Critical Appraisal Skills Programme. (2018). CASP Systematic Review Checklist [Organization]. In *CASP - Critical Appraisal Skills Programme*. https://casp-uk.net/casp-tools-checklists/.

EPPI-Centre. (2003). *Review guidelines for extracting data and quality assessing primary studies in educational research* (Guidelines Version 0.9.7). Social Science Research Unit.

University of Glasgow. (n.d.). *Critical appraisal checklist for a systematic review* [Checklist]. Department of General Practice, University of Glasgow.

Wells, G., Shea, B., O'Connell, D., Robertson, J., Welch, V., Losos, M., & Tugwell, P. (2014). The newcastle-ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. *Ottawa Health Research Institute Web Site*, *7*.

Wu, X., & Zhao, Y. (2021). Degree centrality of a brain network is altered by stereotype threat: Evidences from a resting-state functional magnetic resonance imaging study. *Frontiers in Psychology*, *12*, 705363. https://doi.org/10.3389/fpsyg.2021.705363