Brown and Harkins (2016)

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

☒ Specific focus of this data extraction (please specify)

This data extraction focuses specifically on the aspects of the study related to working memory performance under stereotype threat conditions, as relevant to hypothesis H3 in the preregistration.

Study aim(s) and rationale

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

☒ Explicitly stated (please specify)

The study was explicitly informed by existing research on stereotype threat and working memory. The authors cite numerous previous studies examining the effects of stereotype threat on cognitive performance and working memory capacity.

Do authors report how the study was funded?

☒ Explicitly stated (please specify)

The research was funded by the U.S. Army Research Institute for the Behavioral and Social Sciences (Contract W5J9CQ-12-C-0046).

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

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

The study focuses on stereotype threat and its effects on mind-wandering and working memory performance.

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

The population focus is female university students.

What is the relevant age group?

☒ 17 - 20

☒ 21 and over

The study used undergraduate students, likely spanning ages 17-22+.

What is the sex of the population focus/foci?

☒ Female only

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

☒ Higher education institution

In Which country or cuntries was the study carried out?

☒ Explicitly stated (please specify)

The study was carried out in the United States at Northeastern University.

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

The study examines how stereotype threat affects performance on a sustained attention task (SART) used as an indirect measure of mind-wandering and working memory. It compares performance when the SART is framed as related or unrelated to math ability and gender stereotypes.

What are the study reserach questions and/or hypotheses?

☒ Explicitly stated (please specify)

The study tests two main hypotheses:

1. The mere effort account predicts that describing the SART as related to math/gender stereotypes will eliminate performance differences between stereotype threat and no threat conditions, and lead to better performance than when the SART is described as unrelated.

2. The working memory deficit account predicts that stereotype threat will impair SART performance regardless of how the task is framed.

Methods - Design

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

☒ Explicitly stated (please specify)

The study measures: - Commission errors on the SART - Omission errors on the SART
- Anticipation errors on the SART - Reaction time coefficient of variation on the SART - Manipulation check responses

Study timing

☒ Cross-sectional

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

☒ Only after

Measurements were taken only after the stereotype threat manipulation and SART task framing.

Methods - Groups

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

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

Participants were prospectively allocated to different stereotype threat and task framing conditions.

How do the groups differ?

☒ Explicityly stated (please specify)

The groups differed based on: 1) Whether they received a stereotype threat manipulation for an anticipated math test (threat vs no threat) 2) Whether the SART was framed as related to math/gender stereotypes or as an unrelated attention task

Number of groups

☒ Four or more (please specify)

There were 4 groups in a 2x2 design: 1) GRE stereotype threat / SART stereotype threat 2) GRE stereotype threat / SART unrelated 3) GRE no threat / SART stereotype threat 4) GRE no threat / SART unrelated

Was the assignment of participants to interventions randomised?

☒ Random

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

☒ Not stated/unclear (please specify)

The paper does not provide details on whether allocation was concealed.

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

☒ Yes

Study design summary

This was a 2x2 between-subjects experimental design. The independent variables were stereotype threat condition for an anticipated math test (threat vs no threat) and framing of the SART task (related to math/gender stereotypes vs unrelated attention task). The dependent variables were measures of performance on the SART. Participants were randomly assigned to conditions.

Methods - Sampling strategy

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

☒ Implicit (please specify)

The authors do not explicitly state they are aiming for a representative sample, but the use of female university students implies an attempt to represent the population affected by math-gender stereotypes.

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

☒ Not stated/unclear (please specify)

The sampling frame and identification methods are not clearly stated, but appear to be female undergraduate students at the university.

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

☒ Not stated/unclear (please specify)

The specific selection methods are not described.

Planned sample size

☒ Explicitly stated (please specify)

The authors used G*Power to determine a minimum sample size of 53 to achieve 80% power, but continued data collection to the end of the semester, resulting in 73 participants.

Methods - Recruitment and consent

Which methods are used to recruit people into the study?

☒ Not stated/unclear (please specify)

Recruitment methods are not described.

Were any incentives provided to recruit people into the study?

☒ Explicitly stated (please specify)

Participants received partial course credit for participating.

Was consent sought?

☒ Not stated/unclear (please specify)

The consent process is not described.

Are there any other details relevant to recruitment and consent?

☒ No

Methods - Actual sample

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

☒ Explicitly stated (please specify)

73 female undergraduate students participated.

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

☒ Not stated/unclear (please specify)

This information is not provided.

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

☒ Explicitly stated (please specify)

The participants were from the United States, specifically students at Northeastern University.

What ages are covered by the actual sample?

☒ Not stated/unclear (please specify)

Exact ages are not stated, but the sample consists of undergraduate students.

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

☒ Not stated/unclear (please specify)

Socioeconomic status is not reported.

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

☒ Not stated/unclear (please specify)

Ethnicity information is not provided.

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

☒ Not stated/unclear (please specify)

No information is provided about special educational needs.

Is there any other useful information about the study participants?

☒ Not stated/unclear (please specify)

No additional relevant information about the participants is provided.

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

☒ Unclear (please specify)

Without more details on the sampling frame and achieved sample characteristics, representativeness cannot be determined.

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

☒ Not applicable (not following samples prospectively over time)

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 (not following samples prospectively over time)

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 (not following samples prospectively over time)

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

The main types of data collected were: (a) To define the sample: None reported (b) To measure aspects of the sample: - Commission errors on SART - Omission errors on SART - Anticipation errors on SART
- Reaction time coefficient of variation on SART - Manipulation check questionnaire responses

Which methods were used to collect the data?

☒ Self-completion questionnaire
☒ Psychological test

Details of data collection methods or tool(s).

☒ Explicitly stated (please specify)

The Sustained Attention to Response Task (SART) was used to measure mind-wandering/working memory performance. This involved 240 trials where participants had to respond to frequent non-targets and withhold responses to rare targets. Self-report questionnaires were used for manipulation checks.

Who collected the data?

☒ Researcher

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

☒ Details

The authors do not explicitly discuss reliability of their measures, but they use an established task (SART) that has been used in previous research on mind-wandering.

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

☒ Details

The authors do not explicitly discuss validity, but they cite previous research supporting the use of the SART as an indirect measure of mind-wandering.

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

☒ No (please specify)

The experimenter was aware of the stereotype threat condition when administering the SART.

Where were the data collected?

☒ Unclear/not stated (please specify)

The specific location of data collection is not stated, but presumably at Northeastern University.

Are there other important features of data collection?

☒ Details

No other important features of data collection are mentioned.

Methods - Data analysis

Which methods were used to analyse the data?

☒ Explicitly stated (please specify)

The data were analyzed using linear contrasts in a multivariate ANOVA (MANOVA) with all four SART measures as dependent variables.

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

☒ Details

Multivariate ANOVA (MANOVA) with linear contrasts was used. Univariate ANOVAs were also conducted for each dependent measure. The proportion of variance explained by the contrasts was calculated.

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

☒ Details

The authors state they used linear contrasts to directly test the specific predictions made by the mere effort and working memory deficit accounts.

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

☒ 'Intention to intervene'

Analysis was based on the conditions participants were assigned to.

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

☒ Details

The authors do not explicitly discuss reliability of the data analysis.

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

☒ Details

The authors do not explicitly discuss validity of the data analysis.

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

☒ Details

No strategies to control for confounding variables are described.

Please describe any other important features of the analysis.

☒ Details

The authors calculated the proportion of variance explained by the contrasts testing their hypotheses.

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

☒ Details

No other major analytic issues are apparent.

Results and Conclusions

How are the results of the study presented?

☒ Details

Results are presented in text and in a table showing means and standard deviations for each dependent measure by condition.

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

☒ Details

The mere effort account predictions were supported across all four SART measures:

- When the SART was framed as unrelated, GRE-stereotype threat participants performed worse than GRE-no threat participants.
- Describing the SART as related to math/gender stereotypes eliminated this difference.
- Participants in the SART-stereotype threat conditions outperformed those in the SART-unrelated conditions.

The working memory deficit account predictions were not supported.

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

☒ Yes

The authors report effect sizes (partial eta squared) for the contrasts.

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

☒ No

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

☒ Yes (please specify)

The authors report results for all four SART measures they set out to analyze.

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

☒ No

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

☒ Details

The authors conclude that the results support the mere effort account of stereotype threat effects rather than the working memory deficit account. They argue that poor performance on intervening tasks in stereotype threat studies may reflect reduced motivation rather than cognitive deficits.

Quality of the study - Reporting

Is the context of the study adequately described?

☒ Yes (please specify)

The authors provide adequate context, describing relevant background on stereotype threat research and theories.

Are the aims of the study clearly reported?

☒ Yes (please specify)

The aims to test predictions of mere effort vs working memory deficit accounts are clearly stated.

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

☒ No (please specify)

Details on sample identification and recruitment are lacking.

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

☒ Yes (please specify)

The SART procedure and measures are described in sufficient detail.

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

☒ Yes (please specify)

The linear contrast analysis approach is adequately described.

Is the study replicable from this report?

☒ Yes (please specify)

The key procedures and analyses are described in enough detail to allow replication.

Do the authors avoid selective reporting bias?

☒ Yes (please specify)

Results are reported for all measures and conditions.

Quality of the study - Methods and data

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

☒ No concerns

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

☒ No (please specify)

There is no indication of student/parent involvement in the study design or conduct.

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

☒ Yes (please specify)

The authors justify their approach as a way to directly test competing theoretical accounts.

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

☒ Yes (please specify)

The 2x2 experimental design was appropriate for testing the effects of stereotype threat and task framing.

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?

☒ A little (please specify)

The experimental design helps rule out some alternative explanations, but there may be other factors influencing performance that were not controlled for.

How generalisable are the study results?

☒ Details

Generalizability is limited to female university students in math-related stereotype threat situations. Further research would be needed to extend to other populations or domains.

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

☒ Medium trustworthiness (please specify)

The study uses appropriate methods and analyses to test its hypotheses. However, some limitations in sampling and lack of detail about recruitment reduce overall trustworthiness.

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

⊠ Medium trustworthiness

The authors provide a reasonable justification for their conclusions based on the data. However, alternative explanations are not extensively discussed, which somewhat limits the trustworthiness of the conclusions.

CASE CONTROL STUDIES

This section is not applicable as the study is not a case-control study.

COHORT STUDIES

This section is not applicable as the study is not a cohort study.

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

The study clearly addresses the issue of stereotype threat effects on working memory performance, specifying the population (female university students), intervention (stereotype threat manipulation and task framing), and outcomes (SART performance measures).

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

While this is not a review paper, the authors cite appropriate previous research on stereotype threat and working memory to inform their study design.

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

⊠ Can't tell

This is not applicable as this is a primary research study, not a review.

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.

⊠ Can't tell

Not applicable, as this is not a review study.

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

⊠ Can't tell

Not applicable, as this is not a review study.

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)

The overall results support the mere effort account of stereotype threat effects. When the SART was framed as related to math/gender stereotypes, performance improved regardless of stereotype threat condition for the anticipated math test. This contradicts the working memory deficit account.

How precise are the results?

- Are the results presented with confidence intervals?

⊠ No

The results are presented with effect sizes (partial eta squared) but not confidence intervals.

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
☒ Can't tell

The results may be applicable to female university students in similar cultural contexts, but broader applicability would require further research.

Were all important outcomes considered?

☒ Yes

The study considered multiple measures of SART performance to comprehensively assess working memory and attention.

Are the benefits worth the harms and costs?

• Even if this is not addressed by the review, what do you think?
☒ Yes

The study provides valuable insights into stereotype threat mechanisms with minimal risk to participants. The potential benefits in understanding and potentially mitigating stereotype threat effects likely outweigh the costs of conducting the study.

References

Brown, A. J., & Harkins, S. G. (2016). Threat does not make the mind wander: Reconsidering the effect of stereotype threat on mind-wandering. *Motivation Science*, *2*(2), 85–96. https://doi.org/10.1037/mot0000032