Pennington et al. (2019)

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)

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 linked to existing research on stereotype threat effects on women's mathematical performance. It aimed to elucidate the mechanisms underlying stereotype threat effects by testing the mere effort account against the working memory interference account.

Do authors report how the study was funded?

☒ Not stated/unclear (please specify)

The authors do not report how the study was funded.

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 effects on women's mathematical and visuospatial performance, specifically examining the underlying mechanisms of mere effort vs. working memory interference.

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

The population focus is female university students.

What is the relevant age group?

☒ 21 and over

The mean age of participants was 21 years old.

What is the sex of the population focus/foci?

☒ Female only

The study focused only on female participants.

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

☒ Higher education institution

Participants were recruited from a university in the United Kingdom.

In Which country or cuntries was the study carried out?

☒ Explicitly stated (please specify)

The study was carried out in the United Kingdom.

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

The study examined the effects of stereotype threat on women's performance on an anti-saccade eye-tracking task and a modular arithmetic task. It compared negative and positive stereotype priming conditions to a control condition to test mechanisms of mere effort vs. working memory interference.

What are the study reserach questions and/or hypotheses?

☒ Explicitly stated (please specify)

The study aimed to test two competing hypotheses:

1. The mere effort account predicts that stereotype threat will lead to faster correct saccades and more corrective saccades on the anti-saccade task.

2. The working memory interference account predicts that stereotype threat will lead to slower correct saccades and fewer corrective saccades.

For the math task, it was hypothesized that negative stereotypes would impair performance, especially on difficult problems, while positive stereotypes may facilitate performance on simple problems but impair performance on difficult problems.

Methods - Design

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

☒ Explicitly stated (please specify)

The study measured: - Performance on an anti-saccade eye-tracking task (accuracy, saccadic reaction time) - Performance on a modular arithmetic task (accuracy) - Endorsement of gender stereotypes (manipulation checks)

Study timing

☒ Cross-sectional

The study used a cross-sectional design, measuring outcomes at a single time point.

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 manipulations.

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 randomly assigned to one of three experimental conditions: negative stereotype, positive stereotype, or control.

How do the groups differ?

☒ Explicityly stated (please specify)

The groups differed in the stereotype threat prime they received: - Negative stereotype group: primed with negative stereotypes about women's math ability - Positive stereotype group: primed with positive stereotypes about women's math ability
- Control group: no stereotype priming

Number of groups

☒ Three

There were three experimental groups.

Was the assignment of participants to interventions randomised?

☒ Random

Participants were randomly assigned to the three conditions.

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 specify if 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

All groups completed the same tasks, with only the stereotype prime differing between conditions.

Study design summary

This was a randomized experimental study with three between-subjects conditions (negative stereotype, positive stereotype, control). Participants completed an anti-saccade eye-tracking task and modular arithmetic task after receiving their assigned stereotype prime. The study used a cross-sectional design, measuring outcomes at a single time point after the manipulation.

Methods - Sampling strategy

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

☒ Not stated/unclear (please specify)

The authors do not explicitly state if they are aiming for representative findings. They appear to be testing theoretical mechanisms rather than producing population-representative results.

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 specific sampling frame is not described. Participants were recruited from a university in the UK, but details on how they were identified are not provided.

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 planned sample size was 66 participants, based on a power analysis using the effect size from Jamieson and Harkins (2007).

Methods - Recruitment and consent

Which methods are used to recruit people into the study?

☒ Not stated/unclear (please specify)

The specific recruitment methods are not described.

Were any incentives provided to recruit people into the study?

☒ Explicitly stated (please specify)

Participants received course credits or monetary remuneration for their time.

Was consent sought?

☒ Not stated/unclear (please specify)

The paper does not explicitly state if consent was obtained, though this is likely given the university setting.

Are there any other details relevant to recruitment and consent?

☒ No

No other relevant details are provided.

Methods - Actual sample

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

☒ Explicitly stated (please specify)

60 female participants were included in the final analyses.

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

☒ Not stated/unclear (please specify)

The proportion of selected participants who actually participated is not reported.

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

☒ Explicitly stated (please specify)

Participants were from the United Kingdom.

What ages are covered by the actual sample?

☒ 21 and over

The mean age was 21 years (SD = 5.87).

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

☒ Not stated/unclear (please specify)

Socioeconomic status of participants is not reported.

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

☒ Explicitly stated (please specify)

98.3% of participants were White British.

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?

☒ Explicitly stated (please specify no/s.)

66.7% of participants were Psychology students.

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)

The representativeness of the sample is unclear as details of the sampling frame are not provided.

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)

This was not a longitudinal study.

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)

This was not a longitudinal study.

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)

This was not a longitudinal study.

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: Demographics (age, ethnicity), self-reported math ability and domain identification (b) As findings: Anti-saccade task performance (accuracy, reaction times), modular arithmetic task performance (accuracy), manipulation check responses

Which methods were used to collect the data?

☒ Exams
☒ Clinical test

The anti-saccade task was administered using eye-tracking equipment. The modular arithmetic task was administered via computer.

Details of data collection methods or tool(s).

☒ Explicitly stated (please specify)

Anti-saccade task: Used EyeLink 1000 desktop eye-tracker with sampling rate of 1000 Hz. Modular arithmetic task: Administered with E-Prime experimental software. Manipulation checks: Self-report questions.

Who collected the data?

☒ Not stated/unclear

The paper does not specify who collected the data.

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

☒ Details

The authors used established procedures for the anti-saccade task and modular arithmetic task. They also used manipulation checks to assess the effectiveness of the stereotype threat primes.

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

☒ Details

The authors used established tasks from previous research on stereotype threat and inhibitory control. They also included manipulation checks to validate the stereotype threat primes.

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

⊠ Not stated/unclear (please specify)

The paper does not specify if there was concealment of study allocation from those measuring outcomes.

Where were the data collected?

⊠ Explicitly stated (please specify)

Data were collected at a university in the United Kingdom.

Are there other important features of data collection?

⊠ Details

Participants completed the anti-saccade task and modular arithmetic task in a counterbalanced order.

Methods - Data analysis

Which methods were used to analyse the data?

⊠ Explicitly stated (please specify)

The authors used Analysis of Variance (ANOVA) and Bayesian analyses.

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

⊠ Details

ANOVA, Bonferroni-corrected pairwise comparisons, Bayesian analyses using Bayes factors.

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

⊠ Details

The authors state they used Bayesian analyses in addition to NHST to overcome limitations of statistical power inferences.

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

⊠ Not applicable (not an evaluation study with prospective allocation)

This was not an evaluation study.

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

☒ Details

The authors used both NHST and Bayesian analyses to assess the robustness of their findings.

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

☒ Details

The authors used Bayesian analyses to quantify evidence for the null vs. alternative hypotheses.

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

☒ Details

The authors checked that there were no significant differences between conditions in self-reported math ability or domain identification.

Please describe any other important features of the analysis.

☒ Details

The authors conducted a Bayesian meta-analysis combining data from Experiments 1 and 2 to assess the overall evidence.

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

☒ Details

The authors used Bayes factors to interpret the strength of evidence for null vs. alternative hypotheses.

Results and Conclusions

How are the results of the study presented?

☒ Details

Results are presented through descriptive statistics, ANOVA results, p-values, effect sizes, and Bayes factors.

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

☒ Details

The main findings were: - No significant effects of stereotype threat on anti-saccade task performance - No significant effects of stereotype threat on modular arithmetic task performance - Bayesian analyses provided support for the null hypothesis over the alternative hypotheses

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

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

    The authors reported 95% confidence intervals for mean differences.

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

☒ No

    The reporting of data appears thorough.

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

☒ Yes (please specify)

    The authors report on all key variables specified in their aims, including anti-saccade and modular arithmetic task performance.

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

☒ Yes (please specify)

    The authors state that data and materials are available at: https://osf.io/mdwyv/

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

☒ Details

    The authors conclude that their findings do not support either the mere effort or working memory interference accounts of stereotype threat. They suggest stereotype threat effects may be smaller than previously reported and potentially inflated by publication bias.

Quality of the study - Reporting

Is the context of the study adequately described?

☒ Yes (please specify)

    The authors provide adequate context, describing previous research on stereotype threat and the competing theoretical accounts they aimed to test.

Are the aims of the study clearly reported?

☒ Yes (please specify)

    The aims to test the mere effort vs. working memory interference accounts of stereotype threat 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)

While basic sample characteristics are reported, there is limited information on how participants were identified and recruited.

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

☒ Yes (please specify)

The anti-saccade and modular arithmetic tasks are described in detail.

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

☒ Yes (please specify)

The ANOVA and Bayesian analysis methods are adequately described.

Is the study replicable from this report?

☒ Yes (please specify)

The methods are described in sufficient detail to allow replication.

Do the authors avoid selective reporting bias?

☒ Yes (please specify)

The authors report on all key variables specified in their aims and hypotheses.

Quality of the study - Methods and data

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

☒ No concerns

No ethical concerns are apparent.

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 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 methods based on previous research and the need to test competing theoretical accounts.

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

☒ Yes (please specify)

The experimental design with random assignment to conditions was appropriate for testing causal effects of stereotype threat.

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 lot (please specify)

The random assignment and use of manipulation checks help rule out alternative explanations. The Bayesian analyses also quantify evidence for null vs. alternative hypotheses.

How generalisable are the study results?

☒ Details

The results may have limited generalizability due to the specific sample (primarily White British female university students) and the laboratory setting. However, the findings contribute to the broader literature on stereotype threat mechanisms.

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 can be considered of medium trustworthiness. The experimental design, random assignment, and use of established tasks strengthen its validity. However, the limited sample diversity and lack of detail on recruitment methods slightly reduce its overall trustworthiness.

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

☒ High trustworthiness

The authors provide a balanced interpretation of their findings, considering both their results and the broader literature. They use Bayesian analyses to quantify evidence for null hypotheses and discuss potential limitations and implications.

COHORT STUDIES

[Not applicable - this was not a cohort study]

DOES THIS REVIEW ADDRESS A CLEAR QUESTION?

[Not applicable - this is not a review]

ARE THE RESULTS OF THIS REVIEW VALID?

[Not applicable - this is not a review]

WHAT ARE THE RESULTS?

[Not applicable - this is not a review]

WILL THE RESULTS HELP LOCALLY?

[Not applicable - this is not a review]

References

Pennington, C. R., Litchfield, D., McLatchie, N., & Heim, D. (2019). Stereotype threat may not impact women's inhibitory control or mathematical performance: Providing support for the null hypothesis. *European Journal of Social Psychology*, *49*(4), 717–734. https://doi.org/10.1002/ejsp.2540