

Schmader et al. (2009)

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 Study 3 of the paper, which is most relevant to testing Hypothesis 3 from the preregistration about working memory impairment under stereotype threat.

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 informed by previous research on stereotype threat, including work showing that stereotype threat can lead to increased physiological arousal, anxiety, negative thoughts, and reduced cognitive resources. The authors cite and build on work by researchers such as Steele & Aronson (1995), Schmader & Johns (2003), and others.

Do authors report how the study was funded?

☒ Explicitly stated (please specify)

The research was supported by National Institute of Mental Health Grant 1R01MH071749 awarded to the first author.

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 math performance, examining how individual differences in emotion reappraisal tendencies interact with sympathetic activation to predict performance and self-doubt.

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

The population focus is female undergraduate students preparing to take the Graduate Record Examination (GRE).

What is the relevant age group?

☒ 21 and over

The mean age of participants was 23 years ($SD = 4$).

What is the sex of the population focus/foci?

☒ Female only

The study focused exclusively on female participants.

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

☒ Higher education institution

The study was conducted with undergraduate students at a university.

In Which country or countries was the study carried out?

☒ Not stated/unclear (please specify)

The specific country is not explicitly stated, but based on the authors' affiliations and participant demographics, it appears to have been conducted in the United States.

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

The study examines how individual differences in emotion reappraisal tendencies interact with sympathetic activation (measured via salivary alpha-amylase) to predict women's math performance and feelings of self-doubt on a practice GRE test under conditions of stereotype threat.

What are the study research questions and/or hypotheses?

☒ Explicitly stated (please specify)

The main hypothesis for Study 3 was that for women who experience heightened sympathetic activation when taking a practice GRE test, those low in reappraisal tendencies would perform more poorly on math problems and report more self-doubt about their math performance compared to those high in reappraisal tendencies. For those not experiencing heightened activation, reappraisal tendencies were not expected to relate to math performance.

Methods - Design***Which variables or concepts, if any, does the study aim to measure or examine?***

☒ Explicitly stated (please specify)

The study measured: 1. Math and verbal performance on practice GRE tests 2. Sympathetic activation via salivary alpha-amylase (sAA) 3. Emotion reappraisal tendencies 4. Self-reported feelings of doubt during the tests

Study timing

- ☒ Cross-sectional

The study collected data at a single time point, with participants completing practice GRE tests and providing saliva samples.

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

- ☒ Not applicable (not an evaluation)

This was not an evaluation study with an intervention.

Methods - Groups

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

- ☒ 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)

Comparisons were made based on individual differences in emotion reappraisal tendencies and levels of sympathetic activation.

How do the groups differ?

- ☒ Implicit (please specify)

Groups differed based on levels of emotion reappraisal tendencies (high vs. low) and sympathetic activation (high vs. low), though these were treated as continuous variables in the analyses.

Number of groups

- ☒ Other/unclear (please specify)

The study did not have distinct groups, but rather examined interactions between continuous variables.

Was the assignment of participants to interventions randomised?

- ☒ Not applicable (no prospective allocation)

There was no assignment to interventions in this study.

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 applicable (no prospective allocation)

There was no prospective allocation to groups in this study.

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

☒ Yes

All participants went through the same procedures.

Study design summary

This was a cross-sectional study examining individual differences in emotion reappraisal tendencies and sympathetic activation as predictors of women's math performance and self-doubt under conditions of stereotype threat. Participants completed practice GRE math and verbal tests while providing saliva samples to measure sympathetic activation. The study used moderated regression analyses to test the interactions between reappraisal tendencies and sympathetic activation in predicting performance and self-doubt.

Methods - Sampling strategy

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

☒ Implicit (please specify)

While not explicitly stated, the authors appear to be aiming for findings representative of female undergraduate students preparing for the GRE, particularly in relation to stereotype threat effects on math performance.

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 methods used to identify and recruit participants are not clearly described in the paper.

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 in detail. The paper only mentions that participants were screened for certain health conditions that could affect neuroendocrine concentrations.

Planned sample size

☒ Not stated/unclear (please specify)

The planned sample size is not reported in the paper.

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 in the paper.

Were any incentives provided to recruit people into the study?

- ☒ Explicitly stated (please specify)

Participants were compensated with \$15 and a GRE practice guide.

Was consent sought?

- ☒ Not stated/unclear (please specify)

The paper does not explicitly mention obtaining consent from participants.

Are there any other details relevant to recruitment and consent?

- ☒ No

No additional details about recruitment or consent are provided.

Methods - Actual sample***What was the total number of participants in the study (the actual sample)?***

- ☒ Explicitly stated (please specify)

The final sample consisted of 43 female undergraduates.

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

- ☒ Not stated/unclear (please specify)

The paper does not provide information on how many participants were initially selected or approached compared to the final sample size.

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

- ☒ Implicit (please specify)

While not explicitly stated, based on the authors' affiliations and participant demographics, it appears the sample was from the United States.

What ages are covered by the actual sample?

- ☒ 21 and over

The mean age of participants was 23 years ($SD = 4$).

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

☒ Not stated/unclear (please specify)

The socio-economic status of participants is not reported in the paper.

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

☒ Explicitly stated (please specify)

The sample was 79% Caucasian. The ethnicities of the remaining 21% are not specified.

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 of the participants.

Is there any other useful information about the study participants?

☒ Explicitly stated (please specify no/s.)

All participants were English-speaking female undergraduates who were preparing to take the GRE. They were screened for depression, anxiety, malaria, mononucleosis, or recent surgical procedures that could affect neuroendocrine concentrations.

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

☒ Not applicable (no sampling frame)

There is not enough information provided about the sampling frame or recruitment process to assess representativeness.

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 a cross-sectional study and did not follow samples 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)

This study did not follow samples 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)

This study did not follow samples 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: Demographics (age, gender, ethnicity) (b) To measure aspects of the sample as findings: - Math and verbal performance on practice GRE tests - Salivary alpha-amylase (sAA) as a measure of sympathetic activation - Self-reported emotion reappraisal tendencies - Self-reported feelings of doubt during the tests

Which methods were used to collect the data?

- ☒ Self-completion questionnaire
- ☒ Exams
- ☒ Clinical test

Participants completed questionnaires, took practice GRE tests, and provided saliva samples.

Details of data collection methods or tool(s).

- ☒ Explicitly stated (please specify)
 - Practice GRE tests: Math and verbal subtests from “Practice to Take the GRE General Test” (10th ed.)
 - Reappraisal tendency: Gross and John’s (2003) six-item reappraisal scale
 - Self-doubt: Six items rated on a 7-point scale
 - Salivary alpha-amylase: Collected using the spitting method (Navazesh, 1993) and analyzed using an enzyme kinetic method (α -amylase EPS Sys; Roche Diagnostics)

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 report the internal consistency (Cronbach's alpha) for the reappraisal scale ($\alpha = .89$) and the self-doubt measures ($\alpha = .84$ for quantitative section and .86 for verbal section).

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

☒ Details

The authors used established measures (e.g., practice GRE tests, Gross and John's reappraisal scale) and validated methods for collecting and analyzing salivary alpha-amylase.

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

☒ Not applicable (please say why)

This study did not involve allocation to different conditions or interventions.

Where were the data collected?

☒ Explicitly stated (please specify)

Data were collected in a laboratory setting at the university.

Are there other important features of data collection?

☒ Details

Participants were instructed to refrain from smoking, physical exercise, and food for 2 hours before their participation to control for factors that might affect salivary alpha-amylase levels.

Methods - Data analysis

Which methods were used to analyse the data?

☒ Explicitly stated (please specify)

The main analyses used moderated regression analyses to test the interactions between reappraisal tendencies and sympathetic activation (sAA) in predicting math performance and self-reported doubt.

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

☒ Details

The study used hierarchical regression analyses, including tests of simple slopes and simple effects to probe significant interactions.

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

☒ Details

The authors chose moderated regression analyses to test their hypotheses about how individual differences in reappraisal tendencies would interact with sympathetic activation to predict performance and self-doubt.

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 with prospective allocation.

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

☒ Details

The authors do not explicitly discuss reliability of data analysis, but they use established statistical methods and report standardized regression coefficients and p-values.

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

☒ Details

The authors conducted additional analyses to rule out alternative explanations, such as checking that reappraisal tendencies were not associated with sAA levels.

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

☒ Details

The authors residualized sAA assessments for baseline levels taken on a non-test day to control for individual variation in sAA concentrations.

Please describe any other important features of the analysis.

☒ Details

The authors examined effects separately for math and verbal subtests to show specificity to the stereotype-threatened domain (math).

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

☒ Details

The authors note that they collapsed across a reappraisal manipulation that was ineffective, controlling for it in their analyses.

Results and Conclusions

How are the results of the study presented?

☒ Details

Results are presented through text descriptions of statistical findings, including regression coefficients and p-values. Two figures are also used to illustrate key interactions.

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

☒ Details

The key findings were: 1. For women high in math sAA, reappraisal tendencies predicted better math performance and lower self-reported doubt. 2. For women low in math sAA, reappraisal tendencies were unrelated to performance or doubt. 3. These effects were specific to math performance and were not observed for verbal performance.

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

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

☒ No

The study did not report confidence intervals for the effects.

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

☒ Yes (please specify)

The study does not report effect sizes or confidence intervals, which would have provided more information about the magnitude and precision of the observed effects.

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 mentioned in their hypotheses, including math performance, self-reported doubt, sympathetic activation (sAA), and reappraisal tendencies.

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

☒ No

The paper does not mention where the original data are stored.

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

☒ Details

The authors conclude that individual differences in the tendency to reappraise negative emotions interact with sympathetic activation to predict women's math performance and feelings of self-doubt under stereotype threat. They suggest that positive reappraisal

can do more than just reduce the effects of threat; it can actually fuel better performance when arousal is high.

Quality of the study - Reporting

Is the context of the study adequately described?

☒ Yes (please specify)

The authors provide a thorough background on stereotype threat research and explain the rationale for examining metacognitive processes and emotion reappraisal in this context.

Are the aims of the study clearly reported?

☒ Yes (please specify)

The aims and hypotheses for Study 3 are clearly stated in the introduction to that study.

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 demographic information is provided, there is limited information on how the sample was identified and recruited.

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

☒ Yes (please specify)

The methods for collecting data, including the practice GRE tests, saliva samples, and questionnaires, are described in sufficient detail.

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

☒ Yes (please specify)

The statistical analyses, including moderated regression analyses and simple slopes tests, are adequately described.

Is the study replicable from this report?

☒ No (please specify)

While many aspects of the study are well-described, the lack of detail on recruitment procedures and some specifics of the measures used (e.g., full text of questionnaire items) would make exact replication challenging.

Do the authors avoid selective reporting bias?

☒ Yes (please specify)

The authors report on all variables mentioned in their hypotheses and do not appear to selectively report only significant findings.

Quality of the study - Methods and data***Are there ethical concerns about the way the study was done?***

☒ No concerns

There are no obvious ethical concerns raised by the study procedures.

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

☒ No (please specify)

There is no indication that students or parents were involved in the design or conduct of the study beyond participating.

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

☒ Yes (please specify)

The authors provide a clear rationale for examining emotion reappraisal and sympathetic activation in the context of stereotype threat.

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

☒ Yes (please specify)

The cross-sectional design with individual difference measures was appropriate for examining how reappraisal tendencies interact with sympathetic activation to predict performance under 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 little (please specify)

The study controlled for baseline sAA levels and examined effects separately for math and verbal tests to show specificity. However, the correlational nature of the design limits causal inferences.

How generalisable are the study results?

☒ Details

The results may generalize to female undergraduate students preparing for the GRE, particularly in math-related contexts where stereotype threat is likely. However, the sample was predominantly Caucasian, which may limit generalizability to more diverse populations.

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 established measures and appropriate analyses, but the correlational design and limited information on sampling limit the strength of conclusions that can be drawn.

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

☒ Medium trustworthiness

The authors' conclusions are generally in line with their findings, but they appropriately acknowledge limitations of their work and the need for further research.

References

Schmader, T., Forbes, C. E., Shen Zhang, & Berry Mendes, W. (2009). A metacognitive perspective on the cognitive deficits experienced in intellectually threatening environments. *Personality and Social Psychology Bulletin*, 35(5), 584–596. <https://doi.org/10.1177/0146167208330450>