

Response to Reviewers

Reviewed by Kelsey McCune, 20 Apr 2023 20:11

In my review of this Stage 2 manuscript, I found that the authors were completely consistent with the registered report from Stage 1. The one deviation (5-item rather than 11-item scale), and the failure to meet the preregistered sample size were openly stated and logically explained in the context of their study constraints. I was pleased and impressed with how easy it was to read the manuscript (at both stages, really) and to see the additions in the post-study write up. While I am not in the author's field, the discussion and conclusion points seem well founded based on the results, and present important directions for future research.

The only minor comment I have is for the authors to carefully review the text throughout for spelling and grammar errors arising as a consequence of the changes in verb tense.

Response: Thank you for this kind and helpful appraisal of our paper. We have now done a thorough read through of the paper for spelling and grammar errors.

Reviewed by Neil Lewis, Jr., 08 May 2023 01:08

It was fascinating to see the results of “Evaluating the pedagogical effectiveness of study preregistration in the undergraduate dissertation: A Registered Report” after having reviewed the Stage 1 manuscript a few years ago. Overall, I am quite pleased with the authors’ transparent reporting of their results, and the discussion and interpretation of their findings. The only additional (optional) recommendation I have is for the authors to consider incorporating some of the more recent papers on testing open science practices into their discussion/recommendations for future research.

Like the authors, other metascientists have recently been recommending more careful theorizing about (Gervais, 2021) and evaluation of the effects of open science practices (Buzbas et al., 2023; Suls et al, 2022). It could be beneficial to connect the current results with those broader calls about what is necessary for moving the open science movement forward

References

Buzbas, E. O., Devezer, B., & Baumgaertner, B. (2023). The logical structure of experiments lays the foundation for a theory of reproducibility. *Royal Society Open Science*, 10(3), 221042

Gervais, W. M. (2021). Practical methodological reform needs good theory. *Perspectives on Psychological Science*, 16(4), 827-843.

Suls, J., Rothma, A. J., & Davidson, K. W. (2022). Now is the time to assess the effects of open science practices with randomized control trials. *American Psychologist*, 77(3), 467-475

Response: Thank you for your kind and constructive comments on our paper and for this recommendation. We now include some of these examples in the final conclusion of our paper:

“We hope that this study will contribute to the ongoing reappraisal of open science to progress conversations about the robustness, replicability, and reliability of psychological science. In recent years, there have been productive and important considerations of how to maximise the potential of open science practices (see Gervais et al., 2021; Suls et al., 2022) and the present study contributes to these ongoing metascientific efforts”.

Reviewed by Lisa Spitzer, 24 Apr 2023 13:40

I would like to congratulate the authors for completing their study. As for stage 1, it was a great pleasure for me to review the second part of their Registered Report.

Summary: The authors conducted a study among undergraduate psychology students in the UK to assess if preregistration of the final-year dissertation influences attitudes towards statistics and QRPs, and the perceived understanding of open science. The design followed a 2 (preregistration: yes vs. no) between x 2 (timepoint: before and after dissertation) within subjects design. 52 participants were in the experimental group, 37 in the control group. In contrast to their hypotheses, no effects regarding students' attitudes towards statistics and perception of QRPs were found, however, students who had preregistered had higher perceived understanding of open science at Time 2. Additional exploratory analyses showed that students who preregistered reported higher capability, opportunity, and motivation to do so. Qualitative analyses furthermore gave a more thorough insight into perceived benefits and obstacles of preregistration.

I feel that the study can be recommended after some minor points have been addressed. I have summarised my comments below:

- In some places, the past tense was not yet used consistently enough (e.g., L284 “aim”, L546 “are”, L548 “include”, L580 “are”, L586 “does”). I recommend that the authors go over this again and correct the tense if necessary.

Response: Thank you – we have now gone through and amended these tense typos.

- In addition, “preregistration” was mostly written without a hyphen, but in some places with a hyphen (“pre-registration”). I recommend doing this consistently as well.

Response: This has been amended throughout.

- There is currently no “results” header, which might be a bit confusing. Thus, I recommend adding one before any results are presented.

Response: This has been added.

- I would be interested in exploratory analyses of students who wanted to preregister but then did not. Perhaps it might be interesting to look at their results of capability, opportunity, and motivation?

Response: We agree this may be interesting, but in the interest of keeping the paper within scope (and given our already very long word count) we have decided not to follow this up with exploratory analyses.

- Please report p and np^2 values without zero before the decimal point

Response: This has been resolved throughout.

- Whenever percentages are presented, I recommend presenting the percentages first and n second (when I first read the results section, I thought the numbers indicating the n were part of the percentages), e.g., L853: “($n = 29, 55.8%$)” → “(55.8%, $n = 29$)”

Response: This has been amended.

- Some page and line numbers were incorrect/jumbled

Response: We have attempted to rectify this.

- L413 “inclusion criteria was”: “were”

Response: This has been amended.

- L414 “participants confirmed they met this”: “these”

Response: This has been amended.

- L635 “The same sample of students were”: “was”

Response: This has been amended.

- L667 “uploaded”: I recommend using “uploading” instead of “uploaded” to make it more consistent with point 1 (“creating”).

Response: This has been amended.

- L856-858 “Some students engaged with other Open Science practices in their dissertation, including open materials ($n = 37$), open code ($n = 11$) and open data sharing ($n = 22$).”: Please give percentages.

Response: This has been added.

- L862 “(see Supplementary Information; <https://osf.io/v4fb2>, for our full analysis plan)”: Since you have included this table in the manuscript, please refer to the version in the manuscript instead of the supplementary material.

Response: This has been added.

- L957: I did not understand what “JMMG1” means - is this a mistake perhaps?

Response: This has been removed.

- L8-11 (Discussion): You mention the COM-B model in the conclusions, but I would suggest also mentioning the model here when discussing the results concerning reported capability, opportunity, and motivation.

Response: We have now added this in.

- L9 (Discussion): I recommend deleting the “then” since it might be confusing.

Response: This has been removed.

- L18 (Discussion): Here, the authors refer to a paper by Toth et al. (2021). We recently also surveyed psychological researchers regarding attitudes, motivations, and obstacles regarding preregistration, and came to similar conclusions about perceived obstacles. We also found that supervisors were the biggest influence on students’ decision to preregister or not, which also aligns with the arguments made by the authors. Thus, it might be interesting to refer to our paper as well (“Registered report: Survey on attitudes and experiences regarding preregistration in psychological research”, Spitzer & Mueller, 2023, <https://doi.org/10.1371/journal.pone.0281086>) - but of course this is not a must.

Response: Thank you for sharing this great paper! We now cite it throughout.

L52 (Discussion): delete “too”

Response: This has been removed.

- Table 2: I would recommend avoiding the term "The final planned sample size is therefore 200 participants" as this could lead to misunderstandings (even if the section in which the deviation is described is mentioned afterwards). Alternatively, the final *N* could also be mentioned here.

Response: We have removed this.

- Table 3: Please indicate the scale again.

Response: This has been added.

- Table 4: I recommend using the term “fabrication of data” instead of “falsification of data” because this might be confused with Popper’s Falsification Principle.

Response: This has been changed.

Overall, in my opinion, this Registered Report meets PCI-RR's criteria for stage 2 Registered Reports: It is clear which edits were made to the stage 1 Registered Report, and the hypotheses, as well as the reported methods and procedures align with what was planned a priori. The drawn conclusions are justified given the evidence. Deviations are also described and justified in the paper. The biggest deviation is the smaller sample size of only 89 instead of the targeted 200 participants. We already discussed this risk in the stage 1 Registered Report and the authors had implemented respective countermeasures. I find it very important to clarify to the reader that the non-significant findings are probably due to the low power, which I think the authors do to a sufficient extent. Therefore, in my opinion, the fact that the sample size is smaller than planned is not an obstacle for recommendation.

The methodological rigour of the study is commendable. Additionally, I think the authors have done a good job of describing all deviations and limitations. Overall, I believe this study is an important starting point for further discussions, which I look forward to. I hope that the authors find my comments helpful for revising their manuscript.

Response: Thank you very much for your thoughtful and detailed appraisal of our paper.

Reviewed by anonymous reviewer, 16 May 2023 05:54

This Stage 2 report reflects a major effort to evaluate the impacts of undergrad study pre-registration on statistics and open science attitudes.

My most serious concern with this Stage 2 report is the drastic difference in the planned and achieved sample size. While the Stage 1 proposed to collect of final sample of 200, with 100 participants in each group, the final sample comprised less than half of this planned amount, and only 37 subjects in one group. I appreciate that this study was subject to recruitment and retention issues, and that the study was conducted under time pressure, but this strikes me as a major drawback in the Registered Report context. What is the achieved power, based on the analyzed sample, for the effect size previously proposed at Stage 1? This concerns me both in terms of the reliability of the observed effects, as well as our ability to confidently interpret the null findings.

Response: Thank you for highlighting this. We acknowledged at Stage 1 the risks of sample size with this study and discussed too mitigations that we were taking to avoid this risk. We also, as you request, already discuss in depth the issue of power and report a sensitivity power analysis in the Stage 2 paper (see p. 14 and below). Therefore, while we appreciate that this is a concern, we do explicitly discuss this throughout the paper, detail the power of the study, and take care not to overstate our findings based off this. We also note this as a limitation in the Discussion section.

Based on the lowest cell size ($n = 37$), sensitivity power analyses indicate that we could reliably detect an effect size of $np^2 = .10$ for the Group*Time interaction and pairwise

comparisons of $d \Rightarrow .66$ with 80% statistical power, which was higher than planned. All participants provided informed consent

The introduction and hypotheses match the Stage 1.

Response: Thank you.

The procedures seem to adhere to the Stage 1 plan, with minor deviations (e.g., the use of a 5-point COM-B scale rather than 11). However, I believe readers would benefit from an explicit section for 'deviations from registration' that clearly delineates and explains any deviations, and whether or not they change anything about the results interpretation.

Exploratory analyses are justified and informative.

Response: We detail the sample size deviations throughout this section of the paper and also very explicitly note the deviation within the COM-B scale too: "Note that the 5-point scale is a deviation from our Stage 1 Registered Report, which proposed to use an 11-point Likert scale."

The conclusions are largely justified given the evidence, although at points I think they could adhere a bit more closely to the data. E.g., the discussion states: "Our findings suggest that the process of preregistration can bolster students' understanding of Open Science terminology more broadly, which suggests that this practice may indeed be a useful way of providing an entry point into the wider Open Science conversation." Since the study did not assess understanding Open Science terminology, I think it is more appropriate to state that it may improve their confidence with Open Science concepts. Moreover, since most of the study hypotheses were not met, I think that warrants further discussion of why that might be the case and what implications it has for the utility of preregistration. The discussion still clearly leans in the direction of pursuing widespread adoption and investigation of Open Science practices, rather than concluding that pre-registration experience has no influence on understanding statistical rigor or attitudes toward QRPs (as the data suggest).

Response: Thank you for this suggestion. We have amended our wording in the discussion to more closely align with the data to ensure that we are not overstating our findings.