





Incorporating open research practices into the undergraduate curriculum increases understanding of such practices

A recommendation by [Corina Logan](#)  based on peer reviews by [Neil Lewis, Jr.](#), [Lisa Spitzer](#) , [Kelsey McCune](#) and 1 anonymous reviewer of the STAGE 2 REPORT:

Madeleine Pownall, Charlotte R. Pennington, Emma Norris, Marie Juanchich, David Smaile, Sophie Russell, Debbie Gooch, Thomas Rhys Evans, Sofia Persson, Matthew HC Mak, Loukia Tzavella, Rebecca Monk, Thomas Gough, Christopher SY Benwell, Mahmoud Elsherif, Emily Farran, Thomas Gallagher-Mitchell, Luke T. Kendrick, Julia Bahnmueller, Emily Nordmann, Mirela Zaneva, Katie Gilligan-Lee0, Marina Bazhydai, Andrew Jones, Jemma Sedgmond, Iris Holzleitner, James Reynolds, Jo Moss, Daniel Farrelly, Adam J. Parker, and Kait Clark (2023) Evaluating the Pedagogical Effectiveness of Study Preregistration in the Undergraduate Dissertation. PsyArXiv, ver. 1, peer-reviewed and recommended by Peer Community in Registered Reports.

<https://doi.org/10.31234/osf.io/xg2ah>

Submitted: 26 March 2023, Recommended: 18 August 2023

Cite this recommendation as:

Logan, C. (2023) Incorporating open research practices into the undergraduate curriculum increases understanding of such practices. *Peer Community in Registered Reports*, 100437. [10.24072/pci.rr.100437](https://doi.org/10.24072/pci.rr.100437)

Published: 18 August 2023

Copyright: This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

In a time when open research practices are becoming more widely used to combat questionable research practices (QRPs) in academia, this Registered Report by Pownall and colleagues (2023) empirically investigated the practice of preregistering study plans, which allows us to better understand to what degree such practices increase awareness of QRPs and whether experience with preregistration helps reduce engagement in QRPs. This investigation is timely because results from these kinds of studies are only recently becoming available and the conclusions are providing evidence that open research practices can improve research quality and reliability (e.g., Soderberg et al. 2021, Chambers & Tzavella 2022). The authors crucially focused on the effect of preregistering the undergraduate senior thesis (of psychology students in the UK), which is a key stage in the development of an academic. Pownall and colleagues found that preregistration did not affect attitudes toward QRPs, but it did improve student understanding of open research practices. Using exploratory analyses,

they additionally found that those who preregistered were those students who reported that they had more opportunity, motivation, and greater capability. This shows how important it is to incorporate the teaching of open research practices such that students can increase their capability, motivation, and opportunity to pursue such practices, whether it is preregistration or other practices that are better known to reduce QRPs (such as registered reports; Kryptos et al. 2022). After four rounds of review and revisions, the recommender judged that the manuscript met the Stage 2 criteria and awarded a positive recommendation. **URL to the preregistered Stage 1 protocol:** <https://osf.io/9hjbw> **Level of bias control achieved:** Level 6. *No part of the data or evidence that was used to answer the research question was generated until after IPA.* **List of eligible PCI RR-friendly journals:**

- [Advances in Cognitive Psychology](#)
- [Advances in Methods and Practices in Psychological Science](#)
- [Cambridge Educational Research e-Journal](#)
- [Collabra: Psychology](#)
- [Experimental Psychology](#)
- [F1000Research](#)
- [Meta-Psychology](#)
- [Peer Community Journal](#)
- [PeerJ](#)
- [Royal Society Open Science](#)
- [Studia Psychologica](#)
- [Swiss Psychology Open](#)

References:

1. Chambers C. D. & Tzavella, L. (2022). The past, present, and future of Registered Reports. *Nature Human Behaviour*, 6, 29-42. <https://doi.org/10.1038/s41562-021-01193-7>
2. Kryptos, A. M., Mertens, G., Klugkist, I., & Engelhard, I. M. (2022). Preregistration: Definition, advantages, disadvantages, and how it can help against questionable research practices. In *Avoiding Questionable Research Practices in Applied Psychology* (pp. 343-357). Cham: Springer International Publishing.
3. Pownall, M., Pennington, C. R., Norris, E., Juanchich, M., Smaile, D., Russell, S., Gooch, D., Rhys Evans, T., Persson, S., Mak, M. H. C., Tzavella, L., Monk, R., Gough, T., Benwell, C. S. Y., Elsherif, M., Farran, E., Gallagher-Mitchell, T., Kendrick, L. T., Bahnmüller, J., Nordmann, E., Zaneva, M., Gilligan-Lee, K., Bazhydai, M., Jones, A., Sedgmond, J., Holzleitner, I., Reynolds, J., Moss, J., Farrelly, D., Parker, A. J. & Clark, K. (2023). Evaluating the pedagogical effectiveness of study preregistration in the undergraduate dissertation [Stage 2 Registered Report], acceptance of Version 4 by Peer Community in Registered Reports. <https://psyarxiv.com/xg2ah>
4. Soderberg C. K., Errington T. M., Schiavone S. R., Bottesini J., Thorn F. S., Vazire S., Esterling K. M. & Nosek B. A. (2021) Initial evidence of research quality of registered reports compared with the standard publishing model. *Nature Human Behaviour*, 5, 990-997. <https://doi.org/10.1038/s41562-021-01142-4>

Reviews

Evaluation round #4

DOI or URL of the preprint: <https://osf.io/4jfa7>

Version of the preprint: 1

Authors' reply, 17 August 2023

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Corina Logan](#) , posted 17 August 2023, validated 17 August 2023

Minor revision

Dear authors,

Thank you for your revision and for responding to the comments, which addressed almost all of the remaining points. I found the new exploratory analysis result, that those who planned to preregister, but didn't, rated their capability as lower, really interesting and useful information for teachers to have.

Thank you for providing clarifications on how your Stage 1 and previous Stage 2 version show that the reduced sample size was not a deviation. However, the issue of making it clear which of your results were below or above your power threshold is still there. I think the confusion comes down to two things:

1) "This means that we were only able to detect stronger effects rather than moderate effects, of which none were found." It is unclear whether "of which none were found" indicates that you found no strong effects in your results, therefore you were unable to detect any effects in your study, or whether it indicates that you only had strong effects and no moderate effects were found, which would indicate that you had the power to detect your strong effects. Clarifying this sentence would help resolve the issue.

2) It would make it much clearer if, in the Results section, you added text after the presentation of each result to note whether the effect size was larger or smaller than the threshold for which you were able to detect effects at 80% power.

Once this issue has been resolved, I'm happy to issue IPA. I look forward to reading your next version.

All my best,

Corina

Evaluation round #3

DOI or URL of the preprint: <https://osf.io/h2sg7>

Version of the preprint: 1

Authors' reply, 17 August 2023

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Corina Logan](#) , posted 07 June 2023, validated 07 June 2023

Revise and resubmit

Dear authors,

Thank you very much for your revision and point by point response to the comments made in the previous round. I sent the revised version back to one of the reviewers, but, unfortunately, they do not have time to re-review it. Therefore, I am making the decision based on the information I have and I am asking for a revision to re-address some of the reviewer points in a more comprehensive way. Specifically, I would like you to better address the following comments from Round 2 in order to make the most of this excellent piece of research and add more value.

1) Neil Lewis, Jr. noted that "It could be beneficial to connect the current results with those broader calls about what is necessary for moving the open science movement forward". In your revision, you added only one general sentence that briefly cited other articles. Your results could be more meaningfully connected with the broader literature if you provided more detail and delved more in to the interesting practical and theoretical aspects of these connections.

2) Lisa Spitzer expressed interest 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?" In your response, you declined to conduct the analyses due to "keeping the paper within scope" and a "very long word count". At PCI RR, there are no word limits, and, while you might have a target journal in mind that imposes word limits, your article at PCI RR is independent from this. I err in favor of adding value to the research to get as much as you can out of all of your hard work. While it is entirely your choice about whether you conduct a post-hoc analysis, you might consider whether this would add value to the data you were able to collect and, if you feel it relevant, go ahead with it. To be clear, you don't need to re-address this point in your revision, I just wanted to bring it up in the context of there not being any word limits at PCI RR so you are free to conduct the analysis if you want to.

3) The anonymous reviewer had doubts about your Stage 2 sample size (n=89: 52=experimental group, 37=control group) being much smaller than what was expected (n=200: 100=experimental group, 100=control group) at Stage 1. Here is the reviewer's comment:

"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."

This calls into question whether your Stage 2 meets the review criterion "2C. Whether the authors adhered precisely to the registered study procedures" and, consequently, "2E. Whether the authors' conclusions are justified given the evidence" (https://rr.peercommunityin.org/help/guide_for_recommenders#h_6759646236401613643390905). The much smaller sample size was not discussed with PCI RR as this deviation was starting to unfold during the course of the research. I would like a full justification about 1) whether this small sample size is a deviation from the Stage 1 plan and, if so, 2) explain exactly how it differs from the Stage 1 plan and why you think it is still scientifically valid using the details set out in your power analyses, as well as any other pieces of evidence that can show this. 3) If your small sample size is not a deviation from the Stage 1 plan, please explain exactly why and use details from your Stage 1 power analysis (as well as any other evidence you can bring to bear on the issue) to show why this is the case. A summary of these details should also be included in the article to help readers understand this point because this question will come up for future readers as it already has during the review process. I appreciate that you attempted to address

this comment in your response, however, there isn't enough detail in your response for me to be able to empirically evaluate whether your article meets the above two Stage 2 review criteria.

Additionally, I checked for your data and code and was able to find the data sheet (<https://osf.io/download/zdu8f/>), but I was only able to find the code for the Stage 1 power analysis (<https://osf.io/download/jpmbt/>) and not the code for the Stage 2 Results section. I checked your submission checklist and you state that all analysis code has been provided and is at the following URL: https://osf.io/5qshg/?view_only=. However, I was not able to find the code at this repository. Please provide the remaining code and a direct link to the file at OSF (rather than a general link to the broader project).

I'm looking forward to your response and revision.

All my best,
Corina

Evaluation round #2

DOI or URL of the preprint: <https://osf.io/numr3>

Version of the preprint: 1

Authors' reply, 30 May 2023

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Corina Logan](#) , posted 16 May 2023, validated 16 May 2023

Revision

Dear authors,

I have now received feedback from the same four reviewers from Stage 1, and they have mostly minor comments for you to address in a revision. One reviewer has a larger concern about the much smaller than expected sample size and how it effects the results, and it will be important to make sure you carefully respond to these points and revise your manuscript accordingly.

Once I receive the revision and point by point response to all reviewer comments, I will send it back to a subset of the reviewers.

I look forward to your resubmission.

All my best,
Corina

Reviewed by [Kelsey McCune](#), 20 April 2023

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.

Reviewed by Neil Lewis, Jr., 08 May 2023

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

Reviewed by Lisa Spitzer , 24 April 2023

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.
- 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.
- There is currently no “results” header, which might be a bit confusing. Thus, I recommend adding one before any results are presented.
- 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?
- Please report p and np^2 values without zero before the decimal point

- 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: “($n = 29, 55.8\%$)” ❖ “($55.8\%, n = 29$)”
- Some page and line numbers were incorrect/jumbled
- L413 “inclusion criteria was”: “were”
- L414 “participants confirmed they met this”: “these”
- L635 “The same sample of students were”: “was”
- L667 “uploaded”: I recommend using “uploading” instead of “uploaded” to make it more consistent with point 1 (“creating”).
- 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.
- 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.
- L957: I did not understand what “JMMG1” means - is this a mistake perhaps?
- 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.
- L9 (Discussion): I recommend deleting the “then” since it might be confusing.
- 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.
- L52 (Discussion): delete “too”
- 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.
- Table 3: Please indicate the scale again.
- 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.

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.

All the best,
Lisa Spitzer

Reviewed by anonymous reviewer 1, 16 May 2023

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.

The introduction and hypotheses match the Stage 1.

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.

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).

Evaluation round #1

DOI or URL of the preprint: <https://osf.io/2fvpy>

Version of the preprint: 1

Authors' reply, 04 April 2023

[Download author's reply](#)

Decision by [Corina Logan](#) , posted 27 March 2023, validated 27 March 2023

Revision before peer review

Dear authors,

Thank you very much for your Stage 2 submission of "Evaluating the pedagogical effectiveness of study preregistration in the undergraduate dissertation: A Registered Report". I have read your submission with great interest and I have a few revisions I would like to see before I share this with the reviewers.

My main comment is about how to interpret the results given the smaller sample size where the effect sizes fall below the detectable range. In the Stage 1, it was determined that there would be the ability to detect an effect if the effect size of η^2 was 0.04 or greater (for the two-way interaction between group and time) and a d of 0.40 or greater (for the focal pairwise comparison between preregistration and control groups at time 2). In the Stage 2, the sample size was smaller than originally planned and the sensitivity analysis showed that the detectable effect sizes increased from an η^2 of 0.04 to 0.10 and the d changed from 0.40 to 0.66 or greater. The effect sizes you found (η^2 : 0.001-0.05) appear to be lower than the minimum detectable effect size (note that I was not able to find any d effect sizes reported in the results). Therefore, the probability of your detecting any effects is very low. As such, your findings of no correlations or a correlation should be discussed in terms of your inability to have the power to detect a difference (or not) and not that no difference/a difference was found. I suggest adding a sentence about this each time you report a η^2 effect in the Results section, as well as discussing this specific issue in the Limitations section (note that it is a distinct issue from what is currently discussed in the Limitations). You could add that, with this smaller sample size, you were not able to detect moderate effects, only stronger effects, of which, none were found. It could be useful to translate the effect sizes into actual differences in the measured variables to give readers an idea of what kinds of differences between the groups equate to what kinds of differences in the effect sizes (e.g., there would need to be a difference in 2 points on a 5 point scale for a moderate effect, and a difference of 3 or more points on a 5 point scale for a strong effect).

Please put the study design table back in the main manuscript document and add an optional column on the right stating the outcome. The full tracked changes version of the manuscript needs to be available at the link provided in question 2 of the Report Survey. Please also add line numbers to make commenting during the review process more efficient and clear. I include a few other minor comments below.

Many thanks and all my best,

Corina

Minor comments:

- 1) Perhaps the title should be updated to delete "registered report" because it is now a Stage 2?
- 2) Starting in the Abstract and Intro, you changed the term from "statistics anxiety" to "attitudes toward statistics". I think the former is more informative to the reader about what the term means. However, if you consider competence, value and difficulty as attitudes around statistics, then the broader term is more fitting. If you define "attitudes toward statistics" on your first use, then it would be clearer what the term means.
- 3) Page 14: "we could reliably detect an effect size of $\eta^2 = .10$ for the Group*Time interaction and pairwise comparisons of $d \geq 0.66$ with 80% statistical power" add ", which was higher than planned." per your Stage 1 manuscript.
- 4) Please provide a justification for why you changed the 11 point Likert scale to 5 points, as well as evaluating 3 rather than 6 components (page 19).