Overview

The authors present an interesting replication with detailed supplementary material. I think the rationale is convincing and potentially impactful, and the methods are replicable. I like the approach of using the Qualtrics function to generate responses to demonstrate how you will analyse the data. I was not aware of this before, so this will be useful to demonstrate code on how the eventual data will look.

I’ve split my comments into separate sections below on the manuscript, supplementary material, and reproducibility of OSF files. I have also labelled my comments as to address or suggestions. To address are key points to respond to while suggestions are more stylistic that you can ignore.

Manuscript

1. (to address) abstract / methods – In the abstract, you refer to the sample as British Prolific, but American elsewhere. Please clarify/edit which is the best description.
2. (to address) PCIRR Study Design table – Under rationale for deciding the sensitivity of the test, this entry is quite vague. In [the guidance](https://rr.peercommunityin.org/PCIRegisteredReports/help/guide_for_authors#h_27513965735331613309625021), it focuses on the inferential criteria for supporting the hypotheses or not, so it would be worth outlining your choices for alpha, power etc.
3. (to address) Introduction page 13 - In general I find the introduction informative and focused, but one section that would benefit from elaboration is the causal link extension. Is there precedent on how effective this manipulation can be on empathy?
4. (to address) – Method page 15 – the first sentence ends suddenly when outlining the r = .21 effect. Make sure this is explained fully. As an aside, r = .21 is also close to the lower bound (r = .24) of the confidence interval for the replication r = .36 effect in the OSF files. It might be worth mentioning this as additional justification for the sample size and smallest effect size of interest.
5. (suggestion) – Method page 15 – The extension is ultimately exploratory, but it would be interesting to see a comment on the smallest effect size of interest. For the pairwise comparisons, you would essentially have 95% power for effects of d = 0.33, so you could comment on whether this is useful or not in relation to past research on these kinds of manipulations.
6. (suggestion) – Table 3 page 16 – For someone unfamiliar with McCullough et al., it would be worth adding the geographical origin of their participants to complement specifying your sample will be US American students.
7. (to address) – Table 4 page 17 and measures throughout – The source of the scales is provided for some entries, but not all. For completeness, can you add the source for each entry?
8. (to address) – Measures throughout page 20 onwards – The authors report alpha for the scales which has some limitations as it rests on stringent assumptions (Dunn et al., 2014). Can you either add 95% confidence intervals to the alpha values or consider reporting Omega as Dunn et al. suggest?
9. (to address) – Data analysis page 27 – In the manuscript and supplementary appendix, you explain you will not check for outliers but plan on using Pearson’s r as a parametric test. Parametric tests are generally robust but outliers can be one of the biggest problems (Knief & Forstmeier, 2021), so how do you plan on analysing the data if there are problems with outliers, like switching to a non-parametric or robust test?
10. (to address) – Results page 29 and 34 – The correlations include 16 tests and the ANOVAs includes two DVs and pairwise comparisons. Do you plan on correcting for multiple comparisons and if not, why?
11. (suggestion) – Results Figures 2-9 – In general, I like the ggstatsplots to display the distribution of variables, but the text becomes very small when you have multiple plots like Figures 4-6. They also have information you have not presented or justified in the manuscript like Bayes factors for the correlations, so it might be worth breaking them down and presenting them to the reader where they can easily view them.
12. (to address) – Results page 30 – To address hypothesis 2c, it sounds like you want to compare the magnitude of two correlations. Its not clear how you plan on doing this despite saying they don’t differ significantly for the mock data, so it would be useful to explain your planned statistical methods here like Fisher’s z test.
13. (suggestion) – Table 9 page 32 – For comparing the replication to the original findings, it would be helpful to include the replication effect and 95% CI as an additional column. Otherwise, readers would need to scroll back to the table where you reproduce the results.
14. (to address) – Exploratory analysis page 36 – the final sentence is left as […] whereas the other results are presented with the mock data. It would be worth including the text you plan on including when you have the real data like you did for other sections.

Supplementary appendix

1. (to address) Exclusion criteria page 4 – You say you will focus on the full sample but its not clear how this applies to your general criteria here. Will you analyse all the data first, including those who complete the study multiple times or do not provide consent?
2. (to address) Additional analyses page 6 – This is also relevant to the results and your data analysis plan. You mention equivalence testing here but nowhere in the manuscript, and this section is currently blank. I can see how you plan on interpreting the correlation results in the absence or presence of an effect with Table 9 and LeBel et al. graphs in the supplementary, but for the ANOVAs and mediation results, its not clear whether you are just interested in significant vs non-significant effects, or whether you have a smallest effect size of interest to incorporate into equivalence testing.

OSF

I downloaded all the files from the OSF and I could reproduce all the .Rmd files without edits, so great work here. The materials are also complete with PDF, Word, and Qualtrics file versions available.

In the interests of transparency, I sign my reviews.

Dr James Bartlett, University of Glasgow, UK.

References

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412. <https://doi.org/10.1111/bjop.12046>

Knief, U., & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods*, *53*(6), 2576–2590. <https://doi.org/10.3758/s13428-021-01587-5>