

January 27, 2020

Dear Professor Grüner,

I have now received reports from two reviewers. After re-reading your paper, the comments of the reviewers, and your responses (on behalf of yourself and your co-authors), I am pleased to invite you to submit a revision that addresses the comments below.

1. Reviewer 1 finds your discussion of randomized control trials versus non-randomized control trials to be confusing because it suggests that RCTs do not involve before and after comparisons. Footnote #1 in your paper notes that non-randomized controlled studies (NRSs) may also involve a single measurement per person (which is consistent with the following taxonomy of NRSs: <https://childhoodcancer.cochrane.org/non-randomised-controlled-study-nrs-designs>). However, I don't believe you address Reviewer 1's point that RCTs can also involve before and after measurements. In any case, since both RCTs and NRSs include between- and within-subject designs, it is confusing to equate RCTs with "between-subject" designs, and NRSs with "within-subject" designs. Accordingly, please revise your paper and avoid the use of "between-subject" and "within-subject" designs as proxies for random and non-random assignment of treatment.
2. Reviewer 1 also notes that the issue of external validity is not limited to experimental studies. He/she states that it would be useful to include a discussion of which type of study, experimental or observational, is more likely to be externally invalid. I don't think your response adequately addressed their point. Please revise your paper and expand the discussion of external validity as it differentially applies to experimental and observational studies.
3. Reviewer 2 is unclear why your paper emphasizes p-values, when the arguments apply to frequentist statistical inference in general. In your response, you state "We included an additional note stating that we put a particular focus on p-values because of their high prevalence in empirical research and the scientific criticisms regarding p-value-based statistical practices." I think more than a note is required. The paper should be revised so that p-values are presented as an example of statistical inference, rather than the focus of the study.

I also have some concerns from my own reading of the paper.

4. On page 7 you give the example of a dictator games where "Neither one of the two chance mechanisms – random sampling or randomization – applies. Consequently, there is no role for the *p*-value..." I don't understand that. If the experiment were repeated with the same individuals, their choices could change for any number of reasons. This would introduce sampling error, and thus allow the use of p-values. It seems to me that your argument has more to do with the interpretation of the p-value as opposed to the validity of its usage. Please address this in your revision.
5. Finally, I think your paper would be improved if you tried to imagine under what conditions an observational study would be preferred to an experimental study with respect to statistical inference and the use of p-values. Every observational study can be conceptualized as arising from an experimental design, perhaps a very bad one. On the other side, while experimental

studies can assign treatment randomly in expectation, in any given experiment the assignment might be correlated with confounders. Approaching your subject from this perspective could create more opportunities for you to illustrate your main points. However, this is just a suggestion, not something you are obligated to do.

A revised version of your manuscript that addressed the comments above will be reconsidered for publication. If you choose to submit a revision, be sure to confirm that you made the changes you said you would make in your responses to the reviewers. Also, include a point-by-point response to the comments listed above.

Please note that submitting a revision of your manuscript does not guarantee acceptance. However, I would not ask you to make these changes unless I was optimistic about a positive publication decision.

Once again, thank you for submitting your manuscript to *Economics; The Open-Access, Open-Assessment E-Journal*. I look forward to receiving your revision.

Sincerely,

Bob Reed
Co-Editor, *Economics E-Journal*