Reviewer’s report

Title: Testing an Active Intervention to Deter Researchers’ Use of Questionable Research Practices

Version: 1 Date: 23 Aug 2019

Reviewer: Lex Bouter

Reviewer's report:

This is an interesting randomized intervention study that assesses the effect to a writing exercise of 3-5 minutes (aimed at consistency motives activation versus control) on the defensibility of and the willingness to engage in 15 questionable research practices (QRPs). Although the text is well-written the methods and results sections are difficult to follow due to a number of issues that need to be solved.

It takes the reader a lot of effort to sort out what the primary outcomes are and how these are measured.

− I was confused about the number of main outcomes: two (defensibility and willingness to engage) or five (additionally: impact, rationalization and risk). The abstract and method section made me think two while in the results section five outcomes are presented. Only when consulting S1 and S2 it became clear that the first two mentioned are both primary outcomes and that the additional three are (probably meant to be) secondary outcomes. Please clarify.

− S1 and S2 are also the only source that explain what these five outcome variables are: single items with 7 answer options on a Likert scale. How the five outcomes variables are calculated I could not find. My guess is that the numerical value of defensibility and willingness to engage are the means of the scores of these items over the 15 QRPs presented. And the numerical values of impact, rationalization and risk are probably the mean scores on the 3, 3 and 6 questions that together form these three (secondary?) outcome measures. Please clarify.

− No information is provided on the psychometric properties (reliability, validity, responsiveness, comprehensibility) of these outcome measures, with the exception of some reliability estimates from earlier work (study 1 and 2) on the three outcomes contained in S2 (although it’s unclear for which of the three these Cronbach's alpha's (is that what they are?) refer). Please clarify and expand.

The main message (‘this is a negative RCT’) is buried in an confusing avalanche of subgroup analyses.

− The confusion arises from the fact that the effect of the randomized contrast on the (3 or 5) outcomes is not clearly presented.
The lengthy results section instead presents an enormous amount of subgroup analyses and makes too much fuss about some borderline cases of interaction, many of which have little or nothing to do with the randomized contrast that is the primary determinant studied. Please move all this to a digital supplement. That would make the main text much more clear and substantially shorter.

In fact the RCT is mainly analysed as an observational (non-randomized) study. Please stick to the superior straightforward approach:

- Confounders (other determinants (like gender) other than the primary determinant studied (the contrast of two writing tasks)) are only interesting for adjusting the effects of the primary determinants if needed
- Effect modifiers (interaction), if convincingly present, is only important to decide whether it's necessary to report the overall effect of the primary determinant for subgroups.

I believe that the results should be presented as Mean Differences with corresponding 95% Confidence Intervals. That's much clearer than Means and SD per group followed by 3 statistics. Furthermore, I would recommend to present the results as much as possible in a table and to only highlight the main findings in the text.

Please refrain from statistical spin by labelling non-significant findings as being marginally significant.

Some other major issues are:

- Major study limitations are not acknowledged, e.g.:
  - Attitude is a notorious bad predictor of behaviour (actually engaging in QRPs).
  - The power at subgroup level is (very) low and it's highly likely that a substantial proportion of the reported statistically significant findings are false-positives.
  - The difference between the two writing tasks may be too small and maybe it's a bit unrealistic to expect an effect from 3-5 minutes of writing.
  - The absence of blinding (according to the study protocol in the digital supplements no blinding was attempted) may introduce bias that further diminishes the eventual effect. Maybe a randomized consent design would have been a better alternative.
  - The only limitation you list is not really important: the beauty of RCTs is that the selectivity you mention will hit both randomized groups equally. Consequently it's not a threat to internal validity, although it can compromise extrapolation (external validity or generalizability).

The conclusions drawn (as reported both in the Abstract and the Conclusion section have very little to do with what was actually studied. I strongly recommend to stay closer to your data: no effect was found and it may be that the writing task aimed at consistency motives activation has the unwanted side effect of rationalizing the engagement in QRPs. A comment may be added that this might explain why RCR courses have disappointing little effect and sometimes even an effect in the wrong direction.

I took me a while to understand what you mean by themes 1 - 4. My impression is that you divided the respondents in subgroups based on a typology (the themes) you developed in the qualitative analysis of the texts that resulted from the writing tasks. Under the heading 'Exploratory Narrative Analyses' you present some quantitative results for 2 of these 4 subgroups. I could not find HOW you came to these 4 themes, how many participants ended up in these themes, and why you left out 2 of them in this section. Please clarify and explain and add information or the methods and results of you qualitative data-analysis. You could also consider to move this to the digital supplement or a future article.

Some minor issues are:

- The links to OSF don't work.

- It's rather puzzling why two sources are utilized to recruit participants. Please explain and reflect on the potential differences between these subgroups.

- It's unclear how many were approached to get the 200 participants. Furthermore, no information is given on compliance with the writing task and completion of the survey items. Please add this information, preferably in a flow chart.

- QRP 4 suggests that rounding off a p value from .044 to .4 is wrong, but what you probably mean is then when the p-value is .44 it's wrong to say it's &lt; .4.

**Level of interest**
Please indicate how interesting you found the manuscript:

An article of importance in its field

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable
Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.