**Peer Review and Communication History**

**MS Title**: Do registered reports make scientific findings more believable to the public?

**Author Names**: Elaine Costa, Yoel Inbar, David Tannenbaum

**Submitted: Registered Report stage 1**, June 4, 2021

**Editor First Decision**: Revise & Resubmit

July 19, 2021

Dear Elaine Costa,

I have now received all reviews of your Stage 1 Registered Report submission, “Do registered reports make scientific findings more believable to the public?”, from researchers with relevant expertise. I also independently read the manuscript before consulting these reviews. The reviewers had mostly positive reactions to your manuscript. I agree that your manuscript has important strengths and also that there are some issues that need to be addressed. I therefore encourage you to submit a revised version for further consideration at Collabra: Psychology. You should include a document with a point-by-point response to the reviewers’ comments, outlining each change made in your manuscript or providing a suitable rebuttal.

There are two issues in particular that I would like you to attend to in your revision. First, both Dr. Ebersole and Dr. Chambers commented on a footnote describing an alternative analysis, and I agree with their suggestion: Whatever analysis strategy you propose for the new study, apply this to the preliminary data as well, and report those results in the main manuscript.

Second, one fairly major issue is with regard to the proposed sample size and power analysis. Reviewers felt the sample size was not well justified and would result in an underpowered test based on the preliminary study effect size. I would ask that you plan to increase your sample size, or provide a better justification for the planned sample size.

These are the two main issues that I would like to see addressed in a revision. Reviewers made other good suggestions that you might want to consider in your revision, but I will leave those to your discretion.

In summary, I think this is a promising manuscript and, I hope you will revise it for further consideration at Collabra: Psychology. I look forward to receiving your revision.

Please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained. This is the last opportunity for major editing, therefore please fully check your file prior to resubmission.

If you have any questions or difficulties during this process, please contact us at the editorial office editorialoffice@collabra.org.

We hope you can submit your revision within the next six weeks. If you cannot make this deadline, please let us know as early as possible.

Sincerely,

Laura Scherer

# Reviewer 1

##### Open response questions

### Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

Hi All,

I had the pleasure of reviewing your registered report proposal, “Do registered reports make scientific findings more believable to the public?” In it, you report a pilot study that examines this question and propose a follow-up study. Overall, I think this is a great question to investigate and I like a lot of the study design (especially that random sampling of stimuli). I also love the fact that a study examining if Registered Reports are more believable is being submitted as a Registered Report – that is delightfully meta. I have a few recommendations for additions/changes to the work, none of them major.

The analysis plan for your proposed study makes sense to me, and I think this is a good example of how our understanding of best practices evolve in real time. In a footnote, you mentioned that you applied this new analysis strategy to your pilot study and got similar results. I would add those results to the main text of the paper. They don’t need to be in depth and can just focus on the key comparisons, but I think it will help readers to see the analogous analysis across the studies in text.

I like the revision to the proposed study of dropping “new” from the description of Registered Reports. In case that is indeed strengthening the effect, I’d consider upping the planned sample if you can. If the effect is driven by the idea that new is better, it won’t matter, but if it is enhancing the effect, planning a sample based on what you saw in your pilot may be optimistic. Not a big deal (especially if you do the following point), but it’s something I’d consider.

Given the similarity of the studies, have you considered having planned aggregate analyses? It’s only two studies, but given the sample sizes, I’d add planned analyses for aggregating the results so that you have more precise estimates of the effect of Registered Reports. I imagine you could do this meta-analytically or in a multi-level model framework, nesting within studies.

Finally (and most personal preference-y), I’d consider turning Table 1 into a figure, displaying either the combined data or the data from your proposed study. I would give each scenario a row (like your table) and then plot the means and confidence intervals for each of the four groups. I think that would illustrate your findings well and let readers get a handle on the richness of the data quickly.

Warm Regards,
Charlie Ebersole
University of Virginia

##### Rating scale questions

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  | ✔ |  |  |

# Reviewer 2

##### Open response questions

### Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

This is a signed review (Chris Chambers, Cardiff University). The authors have revised the manuscript thoroughly, converting it into a Stage 1 RR in which an improved design is proposed that includes an active control. I think this is a great improvement and will lead to a more severe and internally valid test of the hypothesis.

My one suggestion for minor revision prior to IPA is that the exact same analysis that is proposed in the new study should be conducted (and reported) on the preliminary data in Stage 1 manuscript. At present, the authors note in Footnote 5 (p9) that “We continue to report the results from our pilot study using cross-random effects models because that is the analysis we pre-registered, but we find qualitatively similar results when using OLS with clustered standard errors.” I agree that it is important to report the preregistered analysis for the preliminary study, but a bridge needs to be formed between the analysis approaches of the two experiments, and “qualitiatively similar” is a little imprecise to connect them securely. I therefore suggest reporting the same regression analysis as an exploratory arm of the preliminary study, perhaps in abbreviated form in the main text of the Stage 1 manuscript, with additional details in a supplement.

##### Rating scale questions

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  | ✔ |  |

# Reviewer 3

##### Open response questions

### Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

Thanks for the opportunity to review (re-review, kind of) the manuscript “Do registered reports make scientific findings more believable to the public?” I really like the idea of making this into a registered report; it is very ‘meta,’ to coin a phrase. I think the idea and experimental design are solid. I also think the tweaks to the experimental design are strong. However, I do have a few concerns I would like the authors to address.

My first issue is that I’m not sure that level of power makes sense, given the pilot results, which themselves were based on a pretty large and well conducted study. The authors say “…a sample size of 1,000 participants provides us with 90% statistical power to detect a significant effect at p < 0.05 for a coefficient size of 0.15 or greater (on a 7-point scale).”

My worry, however, is that I am not sure it is wise to be powering their study for a coefficient of b = 0.15 when the original coefficient was b = 0.09. It is true that b = 0.15 fell inside the original 95% confidence intervals, so we cannot rule it out. Also, if the pilot had not been conducted, and the authors said they were powering for an effect of this size, I would also think that was totally reasonable, and maybe even overkill (as far as I can tell, b = 0.15 is still a very small effect, something like a correlation of r = 0.05 between their credibility rating outcome and a binary indicator of registered report vs non-registered report – a far smaller effect than most psych studies are powered to detect). The problem is the pilot data. In that data, I find that b = 0.15 fell outside of 75% confidence intervals for the original slope. In other words, more likely than not, the new study is powering for an effect larger than we should reasonably presume to exist based on the pilot.

This issue makes me think that this new study is quite likely to result in null results, which the authors state will be taken as inconclusive. But this seems like a wasted opportunity. If possible, I think the authors should perhaps consider collecting a slightly larger sample. As far as I can tell based on some quick simulations, with an N of 1000 and the new fixed effect/clustered SE approach, the authors will have around 66% power to detect an effect the size observed in the original study (b = 0.09). That’s not a disaster, but it’s also not ideal, given that the original slope is our best estimate of what the effect will be in this present study. I extended the N to 1500 and got a power estimate of 83% for the original effect size from the pilot. If it was possible to collect that extra data, that feels like the rigorous thing to do.

In addition, there were a few less important things I wanted to note:

1. I didn’t totally understand the phrase ‘a costly signal’ in the second paragraph of page three. Why are RRs a costly signal? I’m not sure they necessarily involve more work than traditional publishing, it is more that the work of navigating the review process is front-loaded, rather than done after submission (if reviewers require additional data, it can be more time- and resource-consuming to deal with reviewers post-data collection). But perhaps I’m misunderstanding the purpose of the phrase here.
2. I don’t understand the phrase ‘covariances between random effects were fixed at zero’ on page 6. The original model just had random intercepts for participants, and random intercepts for vignettes. How could there be covariance between these random effects?
3. I still can’t escape the feeling that the size of the effect in the pilot is being deliberately obscured. The authors call this effect ‘modest’ in a number of places, and talk about the pilot effect size as a proportion of the slope on the credibility manipulation (0.11), and as a proportion of the standard deviation of the DV (0.06). But ‘modest’ is a subjective term without any clear meaning, and I don’t think this information will be meaningful to readers, and I think it will likely be misunderstood as suggesting the effects were equivalent to, say, an r squared value of 0.06, when in fact the effect was much smaller than that. One simple metric that readers could understand, for example, would be a simple bivariate correlation between the outcome and a dummy indicating registered reports, which in the pilot data was just r = 0.03, a figure often seen as being equivalent to a null result if it occurs in a replication of a well-known finding.

##### Rating scale questions

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  | ✔ |  |

**Author Response**
Aug 2, 2021

See a separate response letter “v2\_RR1\_1462465-cover-letter”

**Editor Final Decision:** Revise & Resubmit

Dec 15, 2021

Dear Elaine Costa,
You have done an excellent job responding to reviewers’ previous comments, and so I am pleased to offer in-principle acceptance (IPA) of your submission at Collabra: Psychology. You may now progress to Stage 2 and complete the study as approved.
Before commencing your study we ask that you:

1. Update the journal office editorialoffice@collabra.org as to the anticipated completion date of your study.
2. Register your approved protocol on the Open Science Framework or other recognized repository, either publicly or privately under embargo until submission of the Stage 2 manuscript. Please note that a time-stamped, independent registration of the protocol is mandatory under journal policy, and manuscripts that do not conform to this requirement cannot be considered at Stage 2. The protocol should be registered unchanged from its current approved state, with the time-stamp preceding implementation of the approved study design. We recommend using the dedicated Stage 1 Registered Report registration mechanism at <https://osf.io/rr/>
Following completion of your study, we invite you to resubmit your paper for peer review as a Stage 2 Registered Report. Please note that your manuscript can still be rejected for publication at Stage 2 if a Collabra: Psychology editor considers any of the following conditions to be met:
• The results were unable to test the authors’ proposed hypotheses by failing to meet the approved outcome-neutral criteria
• The authors altered the Introduction, rationale, or hypotheses, as approved in the Stage 1 submission
• The authors failed to adhere closely to the registered experimental procedures
• Any post hoc (unregistered) analyses were either unjustified, insufficiently caveated, or overly dominant in shaping the authors’ conclusions
• The authors’ conclusions were not justified given the data obtained
We encourage you to read the complete guidelines for authors concerning Stage 2 submissions at <https://drive.google.com/file/d/1mApMxCmFH2GFd3m-qiw89y6rLjUDRDdY/view?usp=sharing>. Please especially note the requirements for data sharing and that withdrawing your manuscript will result in publication of a Withdrawn Registration.
When you are ready to submit your Stage 2 manuscript, please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained.
If you have any questions or difficulties during this process, please contact the editorial office.
I look forward to receiving your Stage 2 manuscript.
Sincerely,
Laura Scherer

**Author Response Registered Report Stage 2**
Nov 19, 2021

See a separate letter “v3\_RR2\_1548688-cover-letter”

**Editor Decision:** Revise & Resubmit

Jan 26, 2022

Dear Dr. Costa,

I took over as action editor of this manuscript because the original associate editor is indisposed. It looks like I have little role to play here, which is a testament to your excellent work! All three of these reviewers were happy to review again here at stage 2, and all three are very happy with the manuscript. So my main job here is to say congratulations, thank you for submitting your work to Collabra, and please attend to the constructive comments of Reviewer 1 (Chris Chambers).

I would like to note, however, that I got the wrong impression of what this study does based on the wording of the manuscript, which perhaps didn’t happen to the other reviewers because they are so familiar with the study. I think the wording of the manuscript will mislead readers because readers will not understand that the study involved a dry description of registered reports without providing any information to the study participants regarding what the motivation of registered reports are. The Introduction of your paper gives a lengthy description of why registered reports were developed and what problems and practices they are intended to combat. But the participants in the study didn’t get any of that, which I think few readers would anticipate from reading your abstract and especially if they begin to read your Introduction and thus are exposed to the aforementioned explanation of why registered reports were developed. After reading about two pages of context and description of actual and potential benefits of registered reports, the manuscript then says “We examine whether non-experts (i.e., members of the general public) are more willing to believe a scientific result published as a registered report”. There is no statement in the Introduction saying that the non-experts would not be getting any of the motivation of registered reports that were extensively described in the previous two pages. I hope you see why I think this will give many readers the wrong impression. The methods section also does not explicitly point out to readers that none of the motivation of registered reports were mentioned to the study participants. So I think only highly alert readers will have this issue in mind. The Discussion also does not mention this issue. In my opinion, the text should be revised to mention this issue in the Introduction and the Discussion. This will, I suspect, result in a substantially higher proportion of readers having a more appropriate grasp of the study. If I were an author, I would also add a sentence to the abstract clarifying what the participants were told about registered reports, namely that the putative benefits are not really explain. However, I do not insist on any of this as your study has already passed multiple rounds of review with the reviewers, and I myself do not appreciate editors trying to get me to change the writing of my paper in such a situation.

I have another comment which is more of a quibble, so feel especially free to ignore this one. The manuscript opens with the statement that new scientific findings are often met with public skepticism, and the examples given are Copernicus’ heliocentric model and Wegener’s theory of continental drift. I don’t think those are good examples because both the heliocentric theory and continental drift theory was highly controversial among researchers for decades, let alone the public. Moreover, they really were new grand theories based on, I believe, bringing together a wide array of old evidence, so I don’t see how such “findings” could have benefited from the registered reports format. A better example might be a more narrow discovery rather than a grand theory, perhaps like the discovery that bacteria cause ulcers. I don’t know the details of Barry Marshall’s papers on the topic, so I’m not positive it’s appropriate, but I do know it was pooh-poohed for decades even though it was the somewhat straightforward conclusion of experiments.

With all the above, I am *not* asking you to change anything substantive about the motivation of the paper or the basic interpretation of the results, that would be inappropriate for an RR, as the motivation and basic interpretation was already approved at stage 1.

About reviewer 1’s comment about p=1.00, I agree that seems weird, but my understanding is that the p-value is technically exactly 1 in a certain sense, but given the quantization of the data (there were only 20 vignettes, so the world can only speak in increments of 1 vignette, meaning that ‘exactly half’ or 10 is best interpreted as somewhere between 9.5 and 10.5, not necessarily as exactly 10), it does seem preferable to report the p-value as >0.99, but it’s up to you.

I second reviewer 1’s suggestion to drop “marginally significant” in favor of “not significant”.

I second reviewer 1’s comment that it is debatable whether combining the pilot with the main study is valid given the confound, so please either drop the corresponding analyses or prominently note this limitation.

Please provide the PDF of the main study survey in ResearchBox, as Reviewer 1 suggests.

Please include the stage 1 manuscript somewhere and link to it in something like the way that Reviewer 1 suggests.

This is the last opportunity for major editing, therefore please fully check your file prior to resubmission.

If you have any questions or difficulties during this process, please contact the editorial office editorialoffice@collabra.org. We hope you can submit your revision within the next six weeks. If you cannot make this deadline, please let us know as early as possible.

best,

Alex Holcombe
Associate Editor, Collabra: Psychology

**Reviewer 1**

**Open response questions**

Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

This is a signed review (Chris Chambers, Cardiff University). I reviewed a previous version of this manuscript and I’m glad to see that the authors pursued my recommended path of using the pilot study as groundwork for a Registered Report that eliminated the confound between novelty bias and RRs in the pilot study. As I suspected might happen, once this confound was removed there was no longer reliable evidence that RRs are associated with a credibility boost among lay readers. I don’t think this is a terribly surprising result. Even after being given a brief tutorial on RRs vs regular publishing, it might be the case that people need a certain level of training or experience in science (including experience of publication bias) to perceive credibility benefits. That said, I do think the study is useful in establishing the boundary conditions of purported benefits of RRs in different settings (not that any judgments about the usefulness of the results are eligible criteria to assess a Stage 2 RR anyway – but I offer this comment simply in passing as someone who is very interested in the impacts of the initiative!)

I think the manuscript is close to being acceptable in its current state. I offer some minor suggestions below for potential revision:

p11: As I understand it, and I may be wrong, I believe p can only rarely (if ever) equal exactly 1, so perhaps replace with p>.99? If this is one of the rare instances where p=1 is appropriate, of course ignore this comment.

p12: “The interaction effect was positive and marginally significant, b = 0.127, SE = 0.074, p = 0.087”. I recommend describing this as simply “non-significant” rather than “marginally significant”. It’s fine to still follow up with the post hoc exploratory tests as described, but I would steer away from “marginal” language in the context of statistical significance because it is only ever used to describe results that are nearly significant, never those that are nearly non-significant, and is therefore inherently prone to interpretative bias.

Given the difference in study design between the pilot and main study (and, in particular, the confound between RR status and novelty in the pilot study), I think it is debatable as to whether the test outcomes reported under “Aggregate Analyses” (p13) that collapse or compare across studies are scientifically valid. The authors also report an exploratory analysis on pp13-14 in which they “regressed credibility scores onto scientific bias beliefs, registered reports, and an interaction term between the two variables, while including study origin fixed effects…” If I understand correctly, this analysis also includes both the pilot and main data, so the same comment applies. I would personally drop these analyses (which I believe were unplanned) as the design difference, in my view, makes the results impossible to interpret securely.

In the researchbox repository, please also provide the survey for the main study in .pdf form, as you did for the pilot study, so that the materials are available in a non-proprietary format.

Finally, in accordance with the Collabra RR guidelines (and general best practice for RRs), I recommend that the authors post the accepted Stage 1 manuscript in the researchbox repository and include a direct URL to the Stage 1 manuscript within the main text of the Stage 2 manuscript (at a prominent place, e.g. end of the Introduction or start of the Method or end of the Abstract etc). I would then suggest adding a couple of sentences along the lines of: “The completed pilot study and main study proposal were reviewed as a Stage 1 Registered Report and achieved in-principle acceptance on [date], prior to data collection for the main study. The Stage 1 manuscript, unchanged from the point of in-principle acceptance, can be downloaded from [URL].”

**Rating scale questions**

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  | ✔ |  |

**Reviewer 2**

**Open response questions**

Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

To the Authors,

First off, sorry for my slow review. I both hope I didn’t hold things up too much and that I am the final holdup. I really enjoyed seeing the results of this project. Overall, I think it’s great. I was particularly interested to see the results concerning beliefs about scientists being biased. I think you did a good job describing those findings and not overclaiming on them. I looked back through my stage 1 review and it looks like you incorporated all my suggestions. I’ve got no further suggestions - both past-me and current-me like this paper. Can’t wait to see it out.

Best,
Charlie Ebersole
American Institutes for Research
cebersole@air.org

**Rating scale questions**

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  | ✔ |  |  |

**Reviewer 3**

**Open response questions**

Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

Great job. You appear to have carried out the study and reported the results faithfully. It may be a null finding, but I think it’s an interesting one.

Also, I just want to say that your materials were so accessible, it took me literally about 5 minutes to download your code and reproduce your analyses. This project is a really impressive demonstration of transparency and rigor. Thank you for your excellent work. I recommend accepting as is.

**Rating scale questions**

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  | ✔ |  |

**Author Response Registered Report stage 2**
Jan 31, 2022

See a separate letter “v4\_RR2\_1561947-cover-letter”

**Editor Final Decision:** Accept

Jan 31, 2022

Dear Dr. Costa,

I am happy to say that your manuscript is now officially accepted for publication in Collabra: Psychology. Congratulations on this excellent work, I think it will make an important contribution to the literature and I look forward to seeing it published! I hope your experiences with Collabra: Psychology have been positive and that you will continue to consider it as an outlet for your work.

As there are no further reviewer revisions to make, you do not have to complete any tasks at this point.

You will be receiving separate correspondence regarding any production and technical comments, data deposits, as well as publication charges. We work with the Copyright Clearance Center to process any applicable APC charges. Please note that your APC transaction must be completed before your article gets published.

You will have an opportunity to check the page proofs before we publish your article. Thank you again for publishing in Collabra: Psychology.

Sincerely,
Alex Holcombe