

February 23, 2023

Dear Julia Schnepf, Gerhard Reese, Susanne Bruckmüller, Maike Braun, Julia Rotzinger, and Sarah E. Martiny,

Thank you for submitting your revised Stage 1 manuscript, “Justice in the Eye of the Beholder: How Comparison Framing Affects the Perception of Global Inequality Through Social Emotions and Justice Sensitivity,” to PCI RR.

I returned your manuscript to the two reviewers who evaluated the first version, and I also read the paper closely myself. We were all in agreement that the revised manuscript is much improved, but that it still requires some revisions before it can be finalized. Accordingly, I am asking that you revise and resubmit your Stage 1 proposal for further evaluation.

Once again, the reviewers have provided thoughtful, detailed comments with which I fully agree, so I urge you to pay close attention to them as you prepare your revision. In my view, the most critical issues are as follows:

1. Reviewer 1 raised important concerns about the inability to reproduce the analyses of the pilot studies. You should examine this carefully to determine what has led to the discrepancy.
2. Both reviewers are still unhappy with your treatment of SDO, and I agree. Given that SDO is clearly having an impact on the results, it would be beneficial to do more with it—either conduct analyses with and without SDO as a control and try to understand why it is impacting the results, or bring SDO into the models as a moderator.
3. I agree with Reviewer 2 that post-hoc power for the pilot studies is not meaningful, and what you should report instead is a sensitivity analysis, i.e., solve for your power given a reasonable effect size or range of effect sizes.
4. I missed this in the first round, but you rely heavily on MANOVAs in your analyses. MANOVAs have the dubious distinction of being misused more often than properly used, as they are only for cases when you are actually interested in the DVs as a multivariate set. That did not seem to be the case here. They should not be used as a “gatekeeper” for univariate tests nor to control error rates. Huberty & Morris (1989) is the classic reference, but this blog post quickly summarizes the issue: <http://psychologicalstatistics.blogspot.com/2021/08/i-will-not-ever-never-run-manova.html?m=1>

When submitting a revision, please provide a cover letter detailing how you have addressed the reviewers’ points.

Thank you for submitting your work to PCI RR, and I look forward to receiving your revised manuscript.

Moin Syed

PCI RR Recommender

## Author Response

Dear Dr. Syed,

thank you very much for your second invitation to revise and resubmit our manuscript entitled “Justice in the Eye of the Beholder: How Comparison Framing Affects the Perception of Global Inequality Through Social Emotions and Justice Sensitivity”.

We have thoroughly revised the manuscript in accordance to the points raised by yourself and the reviewers. More precisely, we re-analyzed our data and now reported one-way ANOVAs or non-parametric tests in order to report the main effects of comparative framing. With regard to the concerns of Reviewer 1, we have re-analyzed all our moderation and moderated mediation models using a z-transformed moderating variable instead of mean centering. In addition, we were able to identify a few typos that explain the replication problem of Reviewer 1. We apologize for these mistakes and now are convinced that our results can be replicated when reanalyzing the data. We have now published the syntax of our analyses together with the data on OSF.

We agree with the suggestion of Reviewer 1 and now have changed the post-hoc power analyses into sensitivity analyses for the results of the preliminary studies, and we have added SDO as an exploratory moderator for the planned main study.

In this response letter, we provide an overview of how we implemented the respective criticism in the revision of our manuscript. Our responses are marked in yellow.

We hope that our revision is in line with your expectations.

---

### Reviews

Reviewed by Mario Gollwitzer, 27 Jan 2023 16:00

The authors have addressed almost all of the issues I had raised in my original review -- nice job! I especially like the more detailed hypotheses regarding the moderating effect(s) of Justice Sensitivity in the preregistered main study (pp. 28-29).

That said, I still have trouble understanding (and, actually, reproducing) the results from the two pilot studies. What caught my attention was that, for instance, in Study 1, the framing x perceived size interaction effects were so highly significant (see Tables 2 and 3), yet the conditional ("simple") effects were not that different from each other, after all...

So I downloaded the raw data from the OSF website and tried to reproduce the results. And that left me with more questions than answers...

One thing I noticed was that the moderator variable ("perceived size"), was heavily skewed in both studies. In Study 1, no single participants chose 1 or 2 (on the 1-7 response scale), while 83% chose 6 or 7. In Study 2, the problem was even larger: Again, no one chose 1 or 2, but 89% chose 6 or 7. So, I doubt whether it makes sense to treat "perceived size" as a continuous moderator here... Dichotomizing this variable might be a solution, but even so, the question is whether it makes sense to treat "perceived size" as a moderator at all if the variance is so small.

→ First, we thank Reviewer 1 for his general acknowledgment of our work on the revision. Second, we thank him for his comment on the distribution of the moderator used in Study 1 and 2. According to Hayes (2018), a deviation from normal distribution should not be a problem for conducting moderation and mediation analyses. Nevertheless, we looked for another way to handle our moderator instead of mean-centering and now used it as a z-standardized variable. In the manuscript, we now

report the new, more robust analyses using the z-standardized moderator. Importantly, these new analyses did not change the results.

Second, I could reproduce most of the findings reported in Tables 2 and 3, but not all of them. For instance, in Study 1, the PROCESS model I ran for the DV "legitimacy" was:

```
PROCESS vars = Treatment Perceived_Difference Legitimacy_scale SDO_scale  
/y=Legitimacy_scale  
/x=Treatment  
/m=Perceived_Difference  
/model=1  
/center=1.
```

One noteworthy difference was the (unconditional) effect of "perceived size", which was  $B = -.19$  in "my" analysis (and  $-.39$  in the authors' analysis; see Table 2). Also, the conditional ("simple") effects of "Treatment" ( $\pm 1$ SD from the sample mean on "perceived size") differ substantially from the numbers reported in the Notes below Table 2. In a similar vein, the results for the DV "Intentions" differ (sometimes only slightly, sometimes more strongly, such as for the "perceived size" effects) from the numbers reported in Table 3. This, I think, needs to be double-checked and clarified, because it has important consequences for the interpretations!

→ We thank Reviewer 1 for looking into the analyses. When reanalyzing the data ourselves, we spotted two typos in our tables. This error was a transcription error while transferring the SPSS output to the Word document. We apologize for this mistake. We have now carefully checked all numbers several times and additionally published the syntax of the new analyses together with the data on the OSF.

Also, I was a bit worried about the fact that controlling for SDO in all of these models obviously had a strong impact on the pattern of findings... In Study 1, for instance, the treatment x perceived size interaction effect on both DVs disappears when SDO is \*not\* included in the respective models... I think this should be discussed more openly and explicitly in the paper.

→ Controlling for SDO mainly had an impact on the findings in the Norwegian sample, but not in the German sample. We will discuss this issue in detail in response to Reviewer 2's comment below.

Finally, I would be careful saying that "The results of the second study mainly replicated the results of the first study..." (p. 21) given that the treatment x perceived size interaction effect on the DV "legitimacy" was not significant in Study 2 (see Table 4).

→ We now used a more careful wording in the discussion of Study 2 (p. 15, p. 21 ff.)

So, all in all, I still do like the preregistered main study! But my doubts about the two pilot studies (and what they can tell us) have actually increased.

**Review PC Registered Report**  
**Justice in the Eye of the Beholder: How Comparison Framing Affects the Perception of Global Inequality Through Social Emotions and Justice Sensitivity**

I enjoyed reading the revised Stage 1 RR and appreciate the authors for taking the time to respond to reviewers. I had a few smaller comments below, most of which deal theoretically with SDO and justice sensitivity.

- I am still not convinced that SDO should be a covariate and not a moderator. The authors write on page 22 that they are interested in “considering individual-level variables that might influence the differential processing of inequality-related information as possible moderators”. That is SDO or system justification/belief in a just world. The justice sensitivity piece doesn’t fit. That being said, if the authors are trying to expand the moderators that have been studied in regard to comparison framing, they should definitely do that! That’s not the vibe I’m getting from the manuscript though.
- The critiques of SDO listed are likely also critiques of Justice Sensitivity (if its been studied yet in such a systematic way as SDO), making the purported contextual nature of SDO as the reason why it needed to stay as a covariate feel less theoretically sound. Indeed, from my read of the articles listed, SDO’s effect becomes weaker or stronger but doesn’t flip in a different direction. Or SDO’s effect is particularly relevant in situations that are in line with the ideology itself (more competitive situations or towards more competitive groups). That doesn’t make SDO contextually dependent in the way the authors seem to be putting forth. Instead, it shows the internally consistent way SDO has been theorized, not giving a strong reason why it shouldn’t be used as a moderator.

→ We disagree to the reviewer’s argument that we should include SDO instead of justice sensitivity as a moderator. As outlined in our last response letter and in our theory section of the registered main study, we perceive justice sensitivity as an appropriate moderator in the context of global inequality because the psychological concept of justice sensitivity should be more directly linked to how framing affects perceptions of legitimacy than SDO does. Further, we argued that SDO has been shown to be a rather poor predictor of the different processing of inequality information (especially in Germany). Furthermore, we checked and found that SDO was not a reliable moderator in the two preliminary studies. Finally, based on previous work in this field (Araújo et al., 2022; Cohrs & Stelzl, 2010; Kunst et al., 2017; and our own experience), we argue that SDO seems to work much better in either highly equal or highly unequal societies (e.g., Norway versus U.S.), but less in medium un/equal societies (such as Germany). Our skepticism towards including SDO as a moderator is not based on the very convincing theoretical framework of SDO. In addition to the problematic issues outlined above, there also seems to be a problem with the translation of the items of the SDO scale into German: Some of our own earlier studies found that the scale consists of items that are not well understood in Germany. This could be the case because some of the wordings do not perfectly fit into the German context. Another reason why SDO may have lower internal consistencies in Germany than, for instance, in the U.S. could be that due to the prevailing sensitivity to the issue of oppression of social groups in Germany, there might be a high social desirability in answering the scale.

However, as also mentioned above in response to Reviewer 1, we are willing to agree on a compromise, and have now included SDO as an exploratory moderator.

- Does SDO interact with the framing manipulation in Studies 1 and 2? The fact that the effect doesn’t work without SDO means that there should be a bit more theorizing as to what exactly does controlling for SDO do/mean. As it stands right now, it feels like controlling for SDO is removing any motivated reasoning based on hierarchy and dominance from the equation, which has implications for which hypothesis (4 or 5) will be more likely to represent the data. That might not be true, but I encourage the authors to think through what this means.

→ We had already addressed this point in the first round of reviews showing that SDO and framing partly interact in the Norwegian sample, but not in the German one. Even though we do not expect a moderating effect of SDO on the influence of the comparative framing, we are now willing to assess SDO as an exploratory variable in our main study.

- Justice sensitivity is mentioned several times in the introduction, but it's not defined in any real way until Study 3.

→ We thank Reviewer 2 for this comment. However, we think that given the structure of our paper and the fact that we mention justice sensitivity in the introduction only when referring to the aim of our third study, which is mainly to test JS as a focal moderator, we would like to stay with our current organization of the paper and let the theoretical part on JS where it actually is.

- While I get the stimuli used the phrase “developing country” instead of “low income”, I believe it's helpful to use the phrase “low income” exclusively in the text. The footnote explains what happened. It is also important because the reason why low-income countries are low income is because of high income countries. Colonialism is the root of the current global inequality, and I believe that's worth mentioning here.

→ We share Reviewer 2's concerns about the term “developing country”. However, unfortunately, we used this term in the stimulus material and we did not want to make post-hoc changes when reporting our material. We completely agree with Reviewer 2 on this point and excluded the terms “developing” and “industrial” countries in the manuscript whenever we did not refer to our materials. Referring to our materials is the only time where we use the term.

- I was a bit confused about the ordering of the items in each study. Why is one dependent variable measured separately from the other, and social emotions measured in between the two dependent variables but they are mediators?

→ We agree with Reviewer 2 that the order of the variables was not fortunate in the preliminary studies. Overall, there is much discussion about whether and how mediators and moderators should be measured in cross-sectional designs (Muller et al., 2011; Preacher, 2015). Overall, we think that the cross-sectional and correlational character of the preliminary studies is a general limitation. This is why we have preregistered a more sophisticated research design in Study 3, in which we will measure the moderator prior to the experiment, and the mediator will be placed at its theoretically presumed position.

- I found Table 1 on page 12 very confusing. Maybe the formatting is off?

→ Table 1 is now deleted and changed into in-text reporting of the results.

- I believe it's APA standard to write the exact  $p$  value unless it is below  $p < .001$ .

→ We now reported all exact  $p$ -values for  $ps > .001$

- The moderated mediation results would be a bit easier to follow in a model format. Something to think about for the final paper.

→ We thank Reviewer 2 for this idea. We will keep it in mind for the later reporting of the preregistered main study.

- Post-hoc power analyses aren't as helpful given they are a monotonic transformation of the  $p$  value. What tends to be more helpful are sensitivity analyses, that shows what is the smallest effect size you could have found with your sample at a given power level (usually 80%).

→ We thank Reviewer 2 for this comment. We have now changed the reported post-hoc power analyses into sensitivity analyses (p. 12, 19).

- The effects in Germany were smaller than the Norwegian sample. Based on the hypotheses, we should have expected larger effects given that Germany has more inequality in it, correct? This is me trying to understand the article's logic.

→ We think that this could be rooted in a similar logic as outlined with regard to the SDO effect. Norway is a country with lower inequality levels, but a context with a high salience of the topic. Whereas Germany is a country context with medium inequality, but the salience of the topic may be lower than in Norway. Another important point is that knowledge about inequality in general (and about undeserved privilege) is a much bigger issue in Norway and figures much more prominently in Norwegian school and university curricula compared to Germany. Thus, the effects of disadvantage framing on social emotions might be stronger in Norway than in Germany because social sensitivity to the topic is higher.

- It took me looking up the justice sensitivity scale to understand why the victim sensitivity subscale wouldn't make sense to use in this context. I thought this scale was measuring the sensitivity of people understanding the perspective of the victim, but that's not it, ha!

→ Yes, we agree that is a somewhat misleading naming of the sub-scale, but that is the label that the authors of the scale introduced.

- The authors discuss the subscale of the justice sensitivity scale was mutually exclusive dynamics, but it is likely that folks will score high (or low) on all three. This is especially true for the beneficiary and perpetrator items. That got me thinking that it might be worthwhile making the scale about global inequality to actually tap into "differential processing of inequality-related information" dimension the authors are interested in. As they write, global inequality can be an abstract concept, making it a bit hazier to understand why people's beliefs about how they deal with inequality in the abstract would be related to their beliefs about global inequality.

- Why is Hypothesis 1 still a hypothesis when it hasn't been supported in Studies 1 and 2?

→ Linguistic framing effects are existent, but typically small (Ansalem & Zoizner, 2022). It is therefore possible that sample sizes in our preliminary studies were simply too small to detect the respective effects. We therefore test Hypotheses 1 and 2 with a bigger sample in Study 3.

## References

Amsalem, E., & Zoizner, A. (2022). Real, but limited: A meta-analytic assessment of framing effects in the political domain. *British Journal of Political Science*, 52(1), 221-237.

<https://doi.org/10.1017/S0007123420000253>

Araújo, R. D. C., Bobowik, M., Vilar, R., Liu, J. H., Gil de Zuniga, H., Kus-Harbord, L., ... & Gouveia, V. V. (2020). Human values and ideological beliefs as predictors of attitudes toward immigrants across 20 countries: The country-level moderating role of threat. *European Journal of Social Psychology*, 50(3), 534-546. <https://doi.org/10.1002/ejsp.2635>

Cohrs, J. C., & Stelzl, M. (2010). How ideological attitudes predict host society members' attitudes toward immigrants: Exploring cross-national differences. *Journal of Social Issues*, 66(4), 673-694. <https://doi.org/10.1111/j.1540-4560.2010.01670.x>

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology*, 89(6), 852–863. <https://doi.org/10.1037/0022-3514.89.6.852>

Kunst, J. R., Fischer, R., Sidanius, J., & Thomsen, L. (2017). Preferences for group dominance track and mediate the effects of macro-level social inequality and violence across societies. *Proceedings of the National Academy of Sciences*, 114(21), 5407-5412. <https://doi.org/10.1073/pnas.1616572114>

Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology*, 66, 825-852. <https://doi.org/10.1146/annurev-psych-010814-015258>