

## **Author response letter to JESP reviews**

### **Reviewer 1**

#### *Introduction*

1. *On p.3, you note that "early conceptions of self-interest were too narrow for overlooking motivations such as being a good group member." Indeed, I wondered how the overall self-interest literature may be overly Westernized, lacking acknowledgment of collectivistic pro-group motivations as self-interest (e.g. an individual acting in favor of the group as it is in their benefit to cooperate given the norms of their society). Tying this into your investigation of communality, I feel your front end could benefit from acknowledging the cultural implications of this investigation (though, I recognize your use of Western samples means you are limited in how much you can truly speak to this)- regardless, perhaps underscoring the need for diversified conceptualizations of self-interest and your initial attempt at unpacking that could increase your manuscript's overall contribution beyond the replication.*

We agree and added these sentences to the introduction: "An important note is that these studies and the current ones were based on Western samples, which leaves a large gap about how most of the world's population would respond to such scenarios. We did not know if the key effect would replicate even within a Western sample."

#### *Study 1*

2. *Are you able to report racial demographic information for your samples?*

No, we do not have this data.

3. *Were no data quality checks used when collecting data from either Prolific or MTurk? See Chmielewski & Kucker (2020) or Kennedy, Clifford, Burleigh and colleagues (2020) for discussion about recent concerns in MTurk data quality specifically. I've heard fewer concerns about Prolific as compared to MTurk, but regardless would love more detail about any efforts to ensure sufficient participant attention levels, survey effort, and overall response quality.*

This text was added to the "Participants" section of the Method: "After every scenario, the participants responded to comprehension questions to make sure they understood the content. Also, the results are overall very similar between the two samples, which suggests there were not major quality issues. Finally, we used planned exclusions for lack of English proficiency or lack of understanding the materials (self-reported)."

4. *I believe readers could benefit from greater explanation of reasoning regarding your final two extension hypotheses (H2E and H2F). While you suggest that communal individuals*

*provide more accurate estimates of self-interest influences as they potentially overestimate self-interest's role to a lesser degree, is it also possible that communal individuals provide inaccurate estimates in the opposite direction (i.e. underestimate the role of self-interest in decision-making of others, believing that people are more motivated by altruism)? As this is the primary extension of this manuscript, I would encourage the authors to provide greater discussion of their rationale for readers to consider.*

Yes, both accounts are possible. We prefer to keep the narrative of the manuscript focused on the primary predictions about estimated vs. actual self-interest. Restructuring the justification of the prediction would have made more sense to us if the findings of this extension had been more revealing.

5. *Were the MTurk and Prolific samples comparatively examined or evaluated in any manner prior to combination to ensure sufficient similarity to warrant grouping? I ask as the difference in paid vs. unpaid blood donations rates for self as listed in Table 4 appear to differ in magnitude between MTurk and Prolific samples- I appreciate you providing some presentation of results broken down across these two samples, but am not sure I understand your reasoning for evaluating some hypotheses across samples and some using the combined sample. In fact, I'm not sure I understand the use of the combined sample at all, as both of your studies individually seem to have robust sample sizes. Greater description of these relevant decisions would benefit readers.*

There appears to be a misunderstanding. Both samples are in the same direction in Table 4. We provided the descriptives for each sample separately and combined, and then for more advanced statistics we used the combined sample. As the results are very similar between the samples, and both populations are similar, and the data and code are openly available, the occasional combination of the samples for simplicity still seems appropriate.

6. *I find this notion of price sensitivity, or sensitivity to payment as predictive of over-estimations of self-interest effects interesting. In tying this to your exploratory age findings (i.e. that younger participants displaying greater self-interest), how does the current study inform self-interest as conceptualized in circumstances of need? Might you anticipate differences in self-interest overestimation across one's lifetime (within-person), or perhaps across SES levels (between-person)?*

We choose not to speculate further on this preliminary, exploratory finding.

#### *Study 4*

7. *Was there a reason the ordinal measure of smoking frequency only used in the MTurk sample, and not the Prolific sample?*

Yes. The different groups worked independently, such that this extension was only in group A. For example in the manuscript, we say: "The replications were conducted in parallel by different teams working independently." These groups chose separate extensions.

8. *How might you suggest future researchers operationalize self-interest given the results of your attempt at a more detailed conceptualization (perhaps to add in your discussion of these findings on p. 25-26)? Given the limited frequency of participants in certain categories of the five examined, might you suggest researchers seeking to continue building on this work revert back to the original measure used in Miller and Ratner (1998), or perhaps compromise with three or four groups to operationalize self-interest? More broadly, I'm trying to understand what was learned via this exercise to contribute to future research, and ensure that contribution is underscored within the manuscript.*

Yes, we have the same goal. It may not be possible to pick operationalizations without knowing the context, sample size, and population being studied. It is also possible to find other real-world scenarios for self-interest like vaccine or mask-wearing. In 1998, smoking was a more appropriate social factor as smoking attitudes were more ambivalent. Currently, there are fewer smokers and attitudes are becoming more negative.

#### *Overall Discussion*

9. *I find the p. 25 explanation of non-effects on four of the eight policies of Study 4 self ratings to not be particularly compelling given your fairly large sample sizes- might there be other reasons for the lack of found effects?*

We suggest this might be an overinterpretation of significance testing. All of the effects are in the same direction, and the hypotheses are not about individual policies.

10. *Acknowledging your lack of support found for a relationship between communality and overestimation of self-interest, what other individual difference characteristics might you suspect relate to one's tendency to over- or underestimate? Further, I'm curious what you might expect had you utilized a more multi-cultural sample with greater representation of Eastern, collectivistic cultures- might your results have emerged as expected if tested in a sample with greater variance in communality?*

We choose not to speculate on these points in the current draft. There is little to go on empirically in our results to inform these.

11. *What are the practical implications of a human tendency to overestimate self-interest? Though an interesting replication, I'm still left a bit with the "why does this matter" question. You do acknowledge that these beliefs "inform not just personal interactions but also public policies and institutions such as the criminal justice system (p. 2)," but I feel your current discussion of*

*practical implications is fairly vague and could benefit from greater grounding, perhaps with concrete examples.*

We made additions to the intro and discussion to give more detail on these policy implications. *Intro:* "As a more specific example, people might believe in 2020 that most people would choose not to wear face masks, choosing self-interest over community disease prevention, but compliance is higher than one might expect." *Discussion:* "Across a society, perceived self-interest is important because it could affect support for laws against individual interests such as on environmental issues. Our results that people overestimate self-interest could potentially help reduce demagoguery appealing to individual interests."

*Overall, I found the study to be methodologically rigorous, though I wonder of the work's theoretical contribution beyond the replication. Though replication is a worthy endeavor in and of itself, I feel the authors have opportunity to either reflect on or contextualize their attempts at extension for greater contribution. I hope these recommendations are helpful, and I wish the authors best of luck as they continue their research!*

We interpret this reviewer to be saying the work is competently done and described, but has doubts about having sufficient contribution for the journal.

*Reviewer #2: The authors provide a replication of people's tendency to overestimate the power of self-interest in determining attitudes. The replications are almost identical to the original paper (Miller & Ratner, 1998) and replicate the original results.*

*I applaud the authors' thorough work in the replication effort and appreciate the importance of replications in psychological science in general.*

*Having said this, it is unclear why this replication will be of interest to JESP readers. As the authors note, results consistent with Miller & Ratner's effect have been found in other papers since the original publication and there is no viable reason to doubt its replicability. Yes, the sample sizes were smaller in the original paper (probably because the cost per participant was higher back then and not because of "lower standards") than the current one and the original paper did not report relevant stats like SDs and variance. However, those facts alone do not justify publication of replication paper. Although replication is important in general, this concern must be balanced with sufficient scientific interest.*

*The addition of communal orientation/motivation does not represent sufficient scientific progress in my view over and above the original Miller & Ratner paper.*

We disagree that the viability of this effect was known prior to these studies, and updated the intro to more clearly explain this. Regardless of previous norms in the field, the power of the original study and the transparency and reproducibility of the report are objectively lower than in the current project. This appears to be less important to the reviewer in judging the likelihood of the original effect than it is for us.

*Moreover, in contrast to the meticulous methodology, I found the introduction and GD to be imprecise in various ways. These parts of the current paper are simply all over the place. The authors seem to oversell the importance of Miller & Ratner's original paper - yes, it's an interesting paper and I personally was inspired by it early in my career - but it's only one of many papers that challenged the influence of self-interest in determining attitudes and behaviors.*

Ok. We made no changes in response to this.

*Moreover, the authors at times bring in citations that are irrelevant to the topic at hand. Among other examples, Epley & Dunning (2000) focus on identifying where people err when overestimating their own moral behaviors and Van Boven et al. (2018) simply note that people can misunderstand political outgroups' attitudes. Neither citation is related to the overestimation of the impact of self-interest.*

This appears to be a mistake by the reviewer. Epley & Dunning (2000) also compared estimated vs. actual preferences of others, which is relevant. Van Boven can be reasonably read to be relevant to self-interest in that members of political groups are motivated to align themselves with those groups in the

classic party-over-policy effect. Perhaps the reviewer thought we were claiming more methodological overlap.

*There is one issue with the authors' methodology that should be explained further (which may have also existed in the original Miller & Ratner paper) - the option "no opinion" in measuring support for policies/donations should not be coded at 0.5 (where 0 = no; 1 = yes). "no opinion" can mean something in-between "yes" and "no" but it also can mean "I genuinely have no opinion" or "I have no information to make this decision." It's not clear that coding these responses as 0.5 is appropriate. I still think the authors will replicate the effects even after excluding the "no opinion" participants, but the analysis should be done both ways.*

We re-ran the analyses excluding the "no opinion" responses and added this to the Results: "As a separate robustness check, the "no opinion" ratings (5.5% of responses) were excluded and these calculations re-run. The result also showed overestimation across both samples (M = 12%) and the interpretation in the Discussion remains consistent with either effect size."

*Finally, in noting the shortcomings of the original Miller & Ratner methodology, the authors should be more precise - they correctly note smaller sample sizes and lack of reporting of some statistics, but they also note "assuredly confirmatory analyses" (p. 6) - why do the authors think Miller & Ratner provided confirmatory analyses? The authors provide basically similar/identical analyses to those used in the original Miller & Ratner. I would delete this point or explain it further.*

These three words were replaced in the revision with "transparency about the timing of analytic decisions".

*In sum, I believe that the introduction and parts of the GD must be more precise and clear, and more realistically place the Miller & Ratner paper in the context of its contribution to research on the influence of self-interest on attitudes.*

*The key difficulty I have in providing a positive recommendation is that while I respect replications I do not believe the current has sufficient scientific value for JESP. I am happy to be contradicted by the other reviewers and AE on this point - this is a matter of opinion on my part, after all.*

Again, we interpret this reviewer to be saying the work is competently done and described, but doesn't meet their novelty or impact thresholds for that journal.