Referee Report

Title: *Revisiting the signal value of emotion in altruistic behavior: Replication and extension Registered Report of Barasch et al. (2014) Studies 3 and 6.*

Authors: Tse Lyn (Rachael) Woo and Gilad Feldman
Corresponding author: Gilad Feldman
Email: gfeldman@hku.hk

Summary
This Stage-1 is a replication and extension of the study 3 and 6 from Barasch et al. (2014). The authors are planning to fully replicate the results from study 3 and 6, with minor deviations, and extend by adding measures of authenticity, rewards motivation, perceived self-focus, and perceived other-focus.

General comments

I would like to thank the authors for undertaking this very interesting replication project and I applaud the choice of using the Registered Report methodology to do so. Overall, I am impressed by the level of detail. All the deviations from the original paper are thoughtfully mentioned and justified. The proposed extensions are carefully designed and I appreciate the authors' goal to build a unified comprehensive comprehensive design. The justification for the sample-size rationale is also well defined and the power analysis well specified. I have some comments that I believe will further improve the clarity of the analyses to be conducted and will improve the statistical precision of the analyses.

Major comments

First, I believe that the authors could improve the hypothesis layout in the PCIRR-Study Design. For example, « Emotions towards victims predict moral character, yet more so when they result in prosocial behavior (compared when they do not) » could be divided into two separate hypotheses for clarity. The same goes for the second hypothesis, which could be divided into multiple hypotheses for clarity. This leads to my second comment.

While reading the stage-1, I found it difficult to distinguish which analyses are part of the confirmatory analysis and which are part of the exploratory analysis. From my understanding, the authors will test the replicability of the hypotheses from the original paper (very well summarised in Table 1). I assume that these will be part of the confirmatory analysis. In addition, the authors will test the inclusion of additional measures (e.g., authentic prosocial motivation), but are these extensions part of the confirmatory analyses or purely exploratory? For example, on page 14, the authors plan to test the following « targets experiencing high distress in the donate condition will be seen as more authentically motivated than targets experiencing low distress in the donate condition. ». However, I fail to see this included in the PCIRR-Study Design Table. The distinction between confirmatory and exploratory analyses in a Registered Report are important, and I believe that the Stage-1 will gain in clarity and statistical precision if the authors provide more detail.

Following my previous comment, I am concerned about the number of hypotheses that the authors will be testing. Replicating and extending findings is good for science to provide more precise effect sizes and a more comprehensive understanding of a phenomenon. However, I am afraid that the extension here will lead to test many hypotheses and increase the likelihood of finding a positive result. If the extensions are part of the confirmatory analysis, I suggest that the authors include a Multiple Hypothesis Testing (MHT) correction procedure. If they are part of the exploratory analysis, I could understand reporting the explorations without but I would still suggest checking the robustness of such findings to a MHT procedure.

The authors plan to replicate using a Prolific sample. Could the authors please clarify the filters that will be applied on Prolific? Will the sample be representative of the US population?

This is not a comment per say but more a question to the authors. I see the benefit of writing a result section with simulated data for the authors. However, I do not see the purpose of including this in the stage-1. In my opinion, a stage-1 should include as much elements as possible up until the result this section. The elements should be described clearly and a referee should be able to « visualize » the result section without seeing it. Adding a simulated data result section, especially with data simulated using random generation in Qualtrics, does not add important elements to the stage-1. If anything, it adds additional work the reviewers as they will have to review this section again for the stage-2. Could the author justify why they decided to include this section?

Minor comments:

This sentence was confusing to read, please consider rephrasing for clarity. « At the time of writing (January 2024), there were 301 Google Scholar citations of the article and have broadly inspired many important follow-up theoretical and empirical articles, such as on the psychology of (in)effective altruism (Berman et al., 2018; Caviola et al., 2020; Caviola et al., 2021), related to the now influential Effective Altruism movement (MacAskill, 2015). »

This fact « Despite its impact, to our knowledge, there are currently no published direct replications of their study. » is already mentioned in text before, please consider deleting.