**Review comments:**

The authors propose to investigate the interaction effect of the offender’s identifiability and the punisher's role on the punishment's severity (this is the main investigation among all the proposed hypotheses). Overall, the proposed research question deserves to be investigated. The study is transparent, and the investigation adds to the literature. However, there are some concerns that the authors need to highlight clearly in their theoretical arguments following up to the hypothesis. My comments:

1. First, I recommend minimising the elaborate presentation of the findings from past research. I understand the importance of emphasising the results of earlier studies, given the effects of identifiability are highly contested. In a potential resubmission, the authors should carefully re-evaluate whether the detailed explanation of the findings is a necessary part of the manuscript. This sometimes diverts the attention from the main arguments for the interaction hypothesis, and the readability suffers. It would be preferable to give this information in a concise tabular format.
2. I had difficulty understanding the distinction between the current inquiry and the study conducted by Kogut (2011). In my opinion, the suggested inquiry appears to be a conceptual replication of Kogut's (2011). If this is indeed the case, it is essential for the authors to explicitly acknowledge this. Given the crucial aspect of the paper is examining the interaction effect between identifiability and the role of the punisher, it is imperative to discuss the need and implications of understanding (or potential revisiting) this effect.
3. The existing literature on the examination of the impacts of identifiability has several mixed findings. Although the authors cited the meta-analysis of Lee and Feeley (2016) as an indication of the identifiability effect, I will be cautious in interpreting the presence of a weak identified victim effect (*r* = .05) advantageous (with the three highest-powered studies in the dataset showing effects that are almost zero. For example, one study: 12,802 participants, r = 0.004). Further, a reanalysis of this meta-analysis by Maier et al. (2023) also uncovered moderate evidence of publication bias (BF01 = 0.11) and strong evidence of the absence of an identifiability effect. In addition, Vu et al. (2024) also found no effect of identifiability on altruism. Given the effect of identifiability is contested so much, I wonder if we need another research paper in this portfolio. – I ask the authors to stress what the investigation might add to the current literature, even if there are boundary effects to it or proves to be no significant difference between any of the proposed conditions in future. This is necessary to interpret in case of different outcomes.
4. Is mediation analysis necessary? The feelings of empathy/moral outrage/blaming toward the offender should be correlated to the extent of punishment based on the identification levels [based on Study 1 of Kogut (2011) or similar analysis from Kogut & Ritov (2005)]. In my opinion, displaying the correlations should reliably predict the directions. In my view, the mediation hypothesis introduces unnecessary complexity in the paper.
5. Furthermore, I failed to comprehend the justification for considering the variables as parallel mediators. It would be great if the authors could illustrate the reasons why the mediator residuals should not be correlated, as this can attenuate the extent of the effects of mediators on the dependent variable (paths b). While the authors referenced the mediation analysis using Preacher & Hayes (2004), it is important to note that collinearity could potentially impact the conclusion, as suggested by Preacher & Hayers (2004) in the same paper.
6. The study design template (page 22) is very difficult to read. The authors should consider making it precise and short. Additionally, I would suggest highlighting any potential post hoc analysis that the authors wish to investigate (for example, ANOVA controlling for empathy).