I really enjoyed reading this Registered Report. It aims to replicate three interesting and influential studies, it was written very clearly, and the methods, as well as the statistical analyses were well chosen. The authors made sure that their Registered Report is very transparent, which I applaud. Overall, I think it will be an important contribution to the field. I do have several comments, which I outline below:

Introduction

1. In the abstract, the authors state: “Our replication [failed to find/found] support for the original findings regarding associations between dispositional emotions and two risk-relevant measures: risk preference and risk optimism [summary effects sizes and CIs].” I suggest that the authors clarify how exactly they will decide whether they failed to find support or found support for the original findings. For instance, how many of the hypotheses need to be (dis)confirmed to state that support was (not) found?
2. The role of framing effects (gain/loss) in this replication is unclear. There is no mention of framing effects in the Introduction, and it only appears in the Methods section. The authors should clarify what is the role of framing effects in the replication.
3. Related to point #2, it would be helpful if the authors could clarify briefly which results presented the original paper they decided **not** to replicate in the current replication, and clarify why that is the case.

Methods

1. The authors state that they decided not to conduct a pretest measuring the ambiguity of events, which might be problematic. A sample of MTurk workers in 2022 quite likely differs in many ways from the original sample of undergraduate students in 2001, which makes it probable that the former group will have a different perception of event ambiguity than the latter. I understand that if the authors decide to conduct a pre-test and then choose the (un)ambiguous events based on its’ results, the final procedure might differ substantially from the original one. However, by not conducting the pre-test, the authors risk that participants’ event ambiguity ratings will be different than in the original study, and as a result, the data might be not usable. Summarizing, I would recommend that the authors conduct a pre-test.
2. In Table 3, the authors state that some descriptive statistics were not reported in the original article. Next, on page 20, the authors state that “The specific items chosen and sued from each scale were not reported in the original article, and we therefore used the scales as is based on reported items in other studies.” I assume that the authors contacted the authors of the original article to clarify these issues, and did not receive a reply (?) I think it would be good if the authors clarified that.
3. The authors state: “Even attentive participants may occasionally answer some attention checks wrong, and we therefore will exclude participants who answered three or more of the five attention checks incorrectly”. Participants who answer three or more attention checks incorrectly are likely very inattentive. Hence, I would suggest that the authors use a more conservative threshold and exclude participants who fail at least one attention check.

Results

1. I suggest that the authors prepare an alternative analysis plan in case their data do not meet the statistical assumptions needed to conduct parametric tests. Such alternative plan could include non-parametric tests or data transformations.
2. I suggest that the authors report replication-related findings without including hope in the regression models, and then report extension-related findings with hope included in the regression models. The reason for that is that hope was not included in the original study, and therefore including it in the replication-related regression models will make the replication (unnecessarily) less alike the original study.
3. In case the authors do not replicate the original findings, they should pre-register analyses that will allow them to conclude that there were indeed no effects, e.g., Bayesian models or equivalence testing.
4. It would be helpful for the reader if the authors clarify in the manuscript which predictors were included in each model.

Thank you for the opportunity to review this interesting Registered Report!

Best regards,

Karolina Ścigała

Postdoctoral researcher, Aarhus University