

Stage 1 Review for *Revisiting the psychological sources of ambiguity avoidance: Replication and extensions of Curley, Yates, and Abrams (1986)* by Sze Ying (Dawn) Yiu, Gilad Feldman

The authors aimed to assess whether ambiguity avoidance could be identified from choice behaviour and whether ambiguity avoidance was distinct from risk avoidance. Further, they aimed to test a range of possible moderator variables of ambiguity attitudes. In general, I thought that this is a nicely constructed study that faithfully follows the target article from which this replication was based, with sensible extensions.

I will structure my review according to the provided RR criteria.

1A. The scientific validity of the research question(s).

The authors' research questions are drawn from a long history of ambiguity attitudes in economics, and the introduction section efficiently summarises this research to demonstrate the scientific justification of their research questions. I do, however, have one point of concern related to convergent/divergent validity:

On Pg 9, para 1 the authors write that "*ambiguity is often conceptualized as an unknown probability of given outcomes.*" This is the economic view of ambiguity (second-order uncertainty) whereas the construct means something slightly different in the psychological literature, where it is linked to attitudes to "stimuli that are complex, unfamiliar, or insoluble" (McLain, 2009, p. 976); in other words, any form of uncertainty. It is sensible to focus on the economic definition since the authors are using the Ellsberg paradigm as their primary measure of ambiguity attitudes. However, it's important to acknowledge that the construct means slightly different things in the two literatures, because if the construct is not defined with precision, results relevant to the economic literature could be unduly extrapolated to the wider psychological literature. (As a side-note, there is some preliminary evidence from my own research that ambiguity attitudes as a personality trait have little apparent relation to ambiguity attitudes as measured with the Ellsberg paradigm; Jach & Smillie, 2019, which has made me particularly cautious to conflate the two. I am interested to see if those findings are replicated in the current study given that you are also measuring trait ambiguity tolerance).

1B. The logic, rationale, and plausibility of the proposed hypotheses, as applicable.

In general, hypotheses are stated precisely and are drawn from the theory outlined in the introduction. They are also mostly drawn from Curley et al., which is appropriate given that this is a close replication.

In my opinion the other-evaluation hypothesis needs more unpacking, especially in the introduction. On page 9 it's written: "[Curley et al.] concluded other-evaluation as the most promising and likely mechanism, in that people avoid ambiguity so that their decision would be justifiable to others given social norms." But why would an ambiguous choice not be justifiable to others given social norms? What particular social norms preclude making an ambiguous choice?

1C. The soundness and feasibility of the methodology and analysis pipeline (including statistical power analysis or alternative sampling plans where applicable).

Broadly this seems very good. I appreciate your power analysis and increasing your sample size to account for the possibility that the original effects were overestimated. I have a few clarifying questions about the method/analysis pipeline:

Page 20: “The mix of chips in Bag 2 Lottery is unknown, but it is likely to be distributed in a biased manner against you.” I’m a bit concerned about the wording of this condition. If participants are told the chips in Lottery Bag 2 have a distribution biased against them, this explicitly gives participants information about the distribution: by definition, this choice then becomes less ambiguous. That suggests that this condition may not genuinely be assessing a moderating variable, but instead reducing the ambiguity and giving participants insight about the most economically wise decision (i.e., to choose the risky option). It would not be surprising if you saw increased ambiguity aversion in this condition. Likewise, if the bias was framed as “friendly” (biased toward you) you might expect to see increased ambiguity seeking, because this provides more information about the economically wise decision. I’m not sure if it is possible to separate a perceived negative bias from a reduction in ambiguity, and that makes me question whether this particular moderating variable is worth including in the analysis. If the authors strongly wish to include it, further justification of its usefulness would be beneficial.

P18: What does “Very low quality choice; 6 = Very high quality choice” here mean? Are participants given a description of what low or high quality means? Without this I feel that participants might find this question itself quite ambiguous to answer.

I appreciate the tables you provided that clearly specify how your replication deviates from the original in terms of hypotheses and methodological details.

Your planned analyses sound appropriate. In addition, some further descriptive details could be useful. For example: looking at mean and variance in choice responses over time (trials); measuring internal consistency for the trials per condition, or some other measure to see how variable participants are in their responses (assuming there is more than 1 trial per condition; see my question about that in Section 1D).

I greatly appreciate the efforts that the authors took to simulate data in the Results section.

Your code looks clear – thank you for extensive commenting.

1D. Whether the clarity and degree of methodological detail is sufficient to closely replicate the proposed study procedures and analysis pipeline and to prevent undisclosed flexibility in the procedures and analyses.

Yes, in general this is clear, with a few exceptions:

“Participants were then randomly assigned to one of the four experimental conditions, answered comprehension and manipulation check questions related to the four conditions

described, presented in random orders.” → it’s unclear whether randomisation is referring to comprehension/manipulation check questions or the conditions.

P19: What are the “funnelling questions”?

Perhaps I missed this, but how many trials per condition are you running in this experiment? I assume more than 1? (If only 1, then I would recommend increasing this to increase precision of your estimates and so that you can estimate reliability (internal consistency)).

Exclusion criteria: There is some ambiguity in the reporting of analysis under exclusion criteria. You have seven points of “general” exclusion criteria, and report that “we will also determine further findings reports with exclusions. In any case, we will report exclusions in detail with results for the full sample and results following exclusions (in either the manuscript or the supplementary).” Does this mean that you will just report one alternative analysis implementing all of these exclusion criteria, or multiple separate analyses, or several analyses grouped in some way? And if you decide to report these additional analyses in the manuscript body, will you still report the full sample results in the manuscript body?

1E. Whether the authors have considered sufficient outcome-neutral conditions (e.g. absence of floor or ceiling effects; positive controls; other quality checks) for ensuring that the obtained results are able to test the stated hypotheses or answer the stated research question(s).

The use of advanced mTurk screening and assessing response quality with exclusion criteria is beneficial here. Additionally, I previously suggested including multiple trials before if the authors hadn’t already. That would help with this criteria point, as you could assess how stable this “overall tendency to avoid ambiguity” (H1) is within a person and condition, not just between-person.

Minor issues/general comments that don’t fit in the above sections:

P9: “Both funds, on average, seem to have the same risk of 50%.” And P22: “Lottery Bag 2 carries the same aggregate risk as Lottery Bag 1.” I know what you mean — that mathematically the ambiguous option is equivalent to 50% risk if you take all the possibilities into account. But readers unfamiliar with the ambiguity literature in economics might find this statement confusing without clarification.

P12: In the table, statistical results for the hypothesis “There is a positive association between risk avoidance and ambiguity avoidance” are not provided.

P17: “Additionally, individual differences measures were added to test exploratory directions.” Even though these tests are exploratory, it could be nice to expand on this section. I see that you are measuring the MSTAT and general risk propensity. It could be worth outlining why/why you would not expect to see a relationship between these and behavioural ambiguity attitudes. For the MSTAT, this relates to convergent validity of the construct across methods, which is an important issue when trying to bridge between disciplines, and relates to the first concern I raised in this review.

References

Jach, H. K., & Smillie, L. D. (2019). To fear or fly to the unknown: Tolerance for ambiguity and Big Five personality traits. *Journal of Research in Personality, 79*, 67–78.

<https://doi.org/10.1016/j.jrp.2019.02.003>

McLain, D. L. (2009). Evidence of the Properties of an Ambiguity Tolerance Measure: The Multiple Stimulus Types Ambiguity Tolerance Scale–II (MSTAT–II). *Psychological Reports, 105*(3), 975–988. <https://doi.org/10.2466/PRO.105.3.975-988>