

## **Reply to decision letter reviews: #175**

We would like to thank the editor and the reviewers for their useful suggestions and below we provide a detailed response as well as a tally of all the changes that were made in the manuscript. For an easier overview of all the changes made, we also provide a summary of changes.

Please note that the editor's and reviewers' comments are in bold while our answers are underneath in normal script.

**A track-changes comparison of the previous submission and the revised submission can be found on: <https://draftable.com/compare/jfXvWYJZYCqU>**

**A track-changes manuscript is provided with the file: "PCIRR-RNR-Curley-et-al-1986-replication&extension-main manuscript-Track-changes.docx"**

### Summary of changes

Below we provide a table with a summary of the main changes to the manuscript and our response to the editor and reviewers:

<b>Section</b>	<b>Actions taken in the current manuscript</b>
General	Ed: We addressed the concerns mentioned by reviewers and implemented revisions based on the feedback provided. R1: We clarified our points on the other-evaluation hypothesis and provided rationale for the hypothesis.
Introduction	Ed: We clarified the concepts of ambiguity in our paper. R1: We distinguished the differences of ambiguity from an economic view and from a psychological perspective, and further elaborated on other-evaluation hypotheses.
Methods	Ed: We responded to methodological concerns mentioned. R1: We further elaborated on our methodology and added two checks; to understand subjects' (1) perspective on the ambiguity of bags and (2) subjects' confidence in the negative bias of ambiguous bag options. We further clarified our objective in closely replicating Curley et al. (1986) and subsequently will not be repeating trials.

---

<b>Section</b>	<b>Actions taken in the current manuscript</b>
Results	R1: We included results of planned exploratory hypotheses in the report.
Discussion	We added points made in the feedback provided to the limitations and future directions section
Reporting	R1: We reworded mentioned paragraphs to reflect clarity in our paper and added statistics from the target's findings

---

*Note.* Ed = Editor, R1/R2 = Reviewer 1/2

## **Response to Editor: Prof. Chris Chamber**

**Two expert reviewers have now evaluated the Stage 1 manuscript. As you will see, both are positive about your submission, which already comes close to meeting the Stage 1 criteria. There are however, a variety of issues to address concerning clarification of concepts, inclusion of additional methodological details (including in vital areas such as exclusion criteria), validity of specific design components, and justification of analytic decisions (including moderator variables). Provided you are able to address all points comprehensively in a revision and response, Stage 1 in-principle acceptance should be forthcoming without requiring further in-depth review.**

Thank you for the reviews, feedback, and the opportunity to revise and resubmit. The feedback has been extremely valuable, and we appreciate your and the reviewers' time and support for our manuscript. Below we respond to and address each of the points made in detail.

## **Response to Reviewer #1: Dr. Hayley Jach**

**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.**

Thank you for the positive opening note and the detailed constructive review.

### **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).**

This is valuable feedback, thank you.

We revised and added further clarifications on the definition of ambiguity and made a brief note that refers to different possible meanings of ambiguity in other literatures. Our definition for this

project is indeed focused on and in alignment with the chosen target for replication Curley et al. (1986). We added this to the introduction in subsection “Ambiguity Avoidance”:

In the psychology literature, ambiguity often refers to “stimuli that are complex, unfamiliar, or insoluble” (McLain 2009, p. 976), yet in the economics and decision-making literature, ambiguity is conceptualized differently, more in line with the target’s reference to “the uncertainty about the success probability itself” (Curley et al, 1986, p. 230). This differentiation was further supported by Jach and Smillie (2019) demonstrating that ambiguity intolerance from the personality perspective has little association to ambiguity aversion from decision theory perspective. For the purpose of this paper we will be focusing on the target’s definition.

We also appreciate the suggested follow up direction about the associations between ambiguity traits and ambiguity attitudes, and we noted this in our plan for the discussion, under section “Limitations and future directions”.

**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 would an ambiguous choice not be justifiable to others given social norms? What particular social norms preclude making an ambiguous choice?**

Good point. Thank you. We agree this has not been made clear enough in the target.

With other-evaluation, there is the added layer of an expectation to justify decisions to others. It seems easier to justify a decision with known probabilities than an outcome with unknown probabilities. We added the following:

People seem to associate more predictable and less ambiguous choices as more accountable, and for social norms to put pressure on members of society to be more predictable, thereby also being more accountable (Lerner & Tetlock, 1999), with conformity meant to reduce anticipate uncertainty and ambiguity to regulate the predictability of the the social environments (Theriault et al., 2021).

**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 be 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.**

We understand the possible issue, and we appreciate the opportunity to improve our explanation.

We added this statement because it does not provide any information about the actual odds and how this bias affects the odds (direction). The probabilities of the mix within the bag can go either way, which makes this provided information irrelevant and nonsensical. It also does not make it less ambiguous, or at least not in a practical sense.

If participants already believe that the ambiguous option was pitted against them, there should not be any differences between the hostility manipulation condition and the regular condition. If we understand your point, then it would seem that what you are suggesting is that there is some ambiguity in whether the ambiguous option is indeed pitted against them or not, and that by

providing them with that information, we are making it clear that it is, thereby removing the ambiguity. This seems like a good point that can be tested empirically. Therefore, we added questions asking participants about their confidence regarding their answers about the ambiguous option being biased against them.

Based on the original paper, and taking into account the procedures used by Curley et al. (1986), they provided the instructions of: “They (participants) were to consider how they would set the composition of Bag 2; and, particularly, if they thought it was “possible to set the composition of Bag 2 so that, no matter what color the player selects, the lottery would be biased against the player” (p.240). They concluded no support for the hostile bias hypothesis.

To address this point, we added two ambiguity questions that would help directly assess whether perception of ambiguity shifted with the hostile manipulation:

How ambiguous is Bag 1/Bag 2 (0 = *Very ambiguous*; 6 = *Very unambiguous*).

We added this to the “Hostility Bias and confidence of possibility of bias (Study 2 manipulation check)” subsection:

Finally, we measured participants’ perception of ambiguity of the bags to examine whether the hostile manipulation shifted participants perceptions of the ambiguity of either bag (0 = *Very ambiguous*; 6 = *Very unambiguous*).

**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.**

Thank you, this is very valuable feedback. We meant to capture justifiability, which the reference to “quality” indeed does not capture well, given that it may embed all kinds of other things that go beyond our aim with this question.

We gave this some thought, and implemented changes in the description.

Instead of

“Very low/high-quality choice”

we changed to

“Not at all/Highly justifiable to others”

(0 = *Not at all justifiable to others*; 6 = *Highly justifiable to others*).

We believe that this change gets us closer to the theory and predictions we were making.

**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).**

Apologies, this might reflect a misunderstanding. We followed the target's paradigm which involved only a single choice, and we do not have repeating trials in this study design.

Future studies can build on what we do here to conduct more elaborate studies with repeating trials.

**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.**

The randomization is in reference to the four experimental conditions. The paragraph has been rephrased to reflect clearer wording. (Methods > Design > Procedure):

Thereafter, all the participants answered all comprehension and manipulation check questions related to the manipulations in the four conditions, and the questions were presented in random order.

**P19: What are the “funneling questions”?**

Thank you. These were provided in the OSF with our Qualtrics survey file and exports.

We now also elaborate that in our methods section:



Exploratory funneling questions asked about participants' seriousness when filling in the survey (1 = Not at all; 5 = Very much), their exposure to similar surveys (Yes/No), their understanding of the purpose of the study (open), and feedback they have for us to improve in our future studies (open).

**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)).**

Our main objective is to follow the original paper closely, and the original study conducted one trial, and therefore we will not be repeating trials. We can see how this might be valuable for future research and therefore added this as a suggestion in our discussion section in the future direction subsection.

**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?**

We will not be doing multiple analyses for each exclusion. We decided to move the exclusions analysis and clarify this in the main manuscript in the methods section under "Exclusion criteria" subsection:

Our reporting will focus on the full sample of participants who completed the study. We predetermined exclusions that we will examine in one additional joint analysis in case we fail to find support for our predictions. Our exclusion criteria is as follows: 1) responding "neutral" for hostility bias manipulation check, pertaining only to Study 2 data analysis, 2) low proficiency of English (self-report < 5, on a 1-7 scale), 3) self-reporting not being serious about filling in the survey (self-report < 4, on a 1-5 scale), 4) self-reporting to have seen or done the survey before. In such a case, we will report and compare both the pre and post exclusion findings and summarize the differences.

**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.**

As we noted above, we will be focusing on a single trial, replicating the target's design and following their data analysis strategy closely.

**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.**

We added clarification of what we mean by mentioning that the ambiguous option is equivalent to a 50% risk. (Introduction > Ambiguity Avoidance):

The second fund is considered equivalent to the first fund with the same risk of 50%, given that the averaging of all the range of possibilities equals 50%.

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

Thank you. We added the corresponding findings for the hypothesis "There is a positive association between risk avoidance and ambiguity avoidance" in Table 1.

**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.**

We are focused on replicating the original paper to revisit and confirm their defined hypotheses. The extensions were meant as exploratory, and therefore we mainly report the associations of these exploratory hypotheses.

Following data collection, we plan to discuss these exploratory directions in the general discussion to interpret these findings and discuss directions for future research (Results > Trait Predictors):

We examined the association between trait risk tolerance and trait ambiguity tolerance with ambiguity avoidance behavior as exploratory directions. We found no support for these associations between risk tolerance and ambiguity avoidance ( $r(998) = .02$ ,  $p = .430$ , 95% CI [-0.04, 0.09]), or for an association between trait ambiguity tolerance and ambiguity avoidance behavior ( $r(998) = .04$ ,  $p = .160$ , 95% CI [-0.02, 0.11]).

We added this to the introduction:

Exploratory extension: Trait predictors

We added two scales of trait risk and ambiguity tolerance to examine associations with ambiguity avoidance behavior. We entertained competing hypotheses detailed as H5 in Table 1.

We left open the possibility of competing hypotheses, now added to Table 1 as H5a versus H5b:

H5a: There is a negative relationship between risk or ambiguity tolerance and ambiguity avoidance. Therefore, the higher tolerance for risk or for ambiguity, the lower the tendency to avoid ambiguity.

H5b (null): There is no relationship between risk or ambiguity tolerance to ambiguity aversion (examined as correlation coefficient confidence intervals lower than  $r = 0.1$ ).

## **Response to Reviewer #2: Dr. Leyla Loued-Khenissi**

### **Overview**

**This study proposes to replicate an experiment from 1986 that provided empirical evidence in support of motivations underlying ambiguity aversion, a putatively “irrational” bias in decision-making. The motivations behind this tendency include “hostility”; anticipated regret; and concerns on post-choice social judgment from others. The replication is important in verifying the conclusions made in the original paper and the authors have clearly done an admirably thorough methodological and statistical groundwork to ensure that their replication minimizes bias. The only criticism of note is in the fuzziness of the term “hostile”, which appears to be a misnomer, but the authors 1) define it in the report; 2) appear to use the original term from the Curley, 1986 paper to name the specific condition.**

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

**The research question posed centers on a replication of a seminal study investigating the causes underlying ambiguity aversion. The validity of the research question, being a replication of an older but foundational study that deserves scrutiny, is sound.**

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

**The proposed study adheres very closely to the original one and in addition, quantifies this adherence (LeBel et al. (2019)). The logic, rationale and plausibility of the hypotheses are therefore applicable, given the aim (replication).**

### **1C,1D,1E**

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

**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.**

**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 methodology of the proposed study is thorough and transparent, including with respect to samples sizes and power estimations, as well as data inclusion/exclusion. The authors included both comprehension and manipulation checks to maximize the reliability of their data. Packages used as well as the code implemented has been made available in addition to an exhaustive explanation of the methodology, making subsequent replication feasible to the community. The level of detail in the report further limits flexibility in the study's subsequent analyses. The authors do propose some exploratory analyses but they appear hypothesis-driven and are not central to the study.**

Thank you very much for the support and encouragement. We are thrilled by this positive feedback.