Dear Dr. Corker,

Thank you for the time and care that you, the reviewers, and the rest of the team at *Advances in Methods and Practices in Psychological Science* have given the manuscript for our project, "Registered Replication Report: A Large Multilab Cross-Cultural Conceptual Replication of Turri, Buckwalter, & Blouw (2015).” We are thrilled to see this project close to publication.

In response to your decision, *Accept with Minor Revisions*, we have prepared a response and revision addressing some of the remaining issues identified by you and the reviewers. The issues that we (you and the lead author team) both consider resolved are not addressed below.

1. From Dr. Corker:

“The inclusion of these variables may have impacted our results” (p. 81) should probably instead say “The omission of these variables may have impacted our results.” Or “Had we been able to examine these variables, we might have observed different results.” Otherwise, this point is addressed.

Our response: Thank you for pointing out this confusing sentence. We have changed the word “inclusion” to “omission” as suggested.

1. From Dr. Corker:

“Given the length and complexity of the paper, together with changes to the text, it is almost impossible to verify whether all deviations from the Stage 1 protocol have been disclosed. As an example of a deviation discovered, consider the issue of “CREP completion certificates.” These certificates are described in the proposal as “a completion certification award for meeting the required sample size” and “presented to participating lab members for their high-quality work upon completion.” They are described in the current version of the paper as “certificates for completing all pedagogical tasks (e.g., site-level analyses).” The authors opt to include “all usable data” in the main analyses, ignoring the certificate. They discuss this decision in the paper and in their response to review (response to me; also response to reviewer 4, point 8). As requested, the authors have presented two versions of the analysis so that readers can transparently see whether this change in the planned sample is impactful. I think this particular issue is resolved, but I raise it as an example of the kinds of changes that have gone on – most of which are not going to be easily discernible by readers. I am conflicted about what, if anything, to do about this issue. As regards my earlier point 3 (changes to analysis strategy), given the inclusion of the robustness tests in the supplemental material, I consider this point resolved.”

Our response: We appreciate your determination that this point is resolved. We have explained or footnoted all deviations throughout the manuscript and hope that this is sufficient for readers to identify them.

In our revision, we have clarified the language around the completion certificates on page 22 and included a footnote to explain the deviation; initially, we planned on applying the standard CREP rules for completion certificates – “post-data collection review of the team’s updated OSF page including anonymized data, a brief writeup of the results with site-level analysis, and a completion pledge.” As discussed in the footnote, we ended up withdrawing the results writeup as a requirement because CREP projects typically require less sophisticated site-level analyses than this project did. During this time, the CREP team was having internal conversations about the fact that many of our published studies included data from projects that did not receive completion certificates but were still considered high-quality because of their preregistrations, pre-data collection review, and other safeguards. In the manuscript, we also clarify that the sample size requirement was reduced because of issues related to COVID, although requirements in general have evolved over the years, particularly during the CREP leadership transition that occurred during active data collection for this study.

1. From Dr. Corker:

“As you will see in the reviews, two of the reviewers did not believe that the current draft emphasized similarities with the original results enough. I am inclined to agree with the reviewer that a statement like “This result did not correspond to that found by Turri et al., who failed to detect a difference in knowledge attribution between these two conditions” is somewhat misleading (because their failure to detect an effect was seemingly due to low power, given consistency in effect size between their result and yours). However, I think we can let these matters of disagreement stand as they are. I’m expressing my disagreement with your interpretation here, and future academic debate might be spurred as a result of your stance. Please note that one of the reviewers has expressed the desire to prepare a commentary, and we are going to allow them to do so, if they so choose. So, this point is resolved, for now.”

Our response: We appreciate your determination that this point is resolved. However, we did want to respond to this comment and make some minor revisions in light of remaining reviewer concerns.

Because we already make the point about the effect size similarity in the abstract and elsewhere in the manuscript, we believe we have sufficiently acknowledged the nuance of our results in general. However, we do acknowledge that the strength and phrasing of a few statements in the prior version may have not been entirely appropriate. We have changed a few sentences in the manuscript accordingly.

In context, we believe the sentence you identified, specifically, is appropriate, however. Our observation of a significant difference between the gettier and knowledge condition does not correspond to Turri et al.’s failure to detect that difference. That sentence is followed by an acknowledgment that “the Darrel vignette from the original study produced the smallest effect size, one similar to that we calculated using the non-significant result from the original study”, thus clarifying that the lack of correspondence referred to the detection of the effect rather than its size.

We appreciate you letting us know about the potential paper commentary. While we hope that some of the revisions we have made will assuage that reviewer’s concerns, we understand that the same results can yield different interpretations. We hope you will give us an opportunity to review and respond to the commentary if we choose to do so.

Comments from Reviewers

1. Reviewer 1: “[The difference in results] is explained on page 69, where the authors state that they have disproven Turri et al.’s claim that “a salient but failed threat to the truth of a judgment does not significantly affect whether it is viewed as knowledge” and that their “results best align with those of other researchers who have found similar effects and concluded that protagonists with luckily true beliefs are less likely to elicit knowledge attributions than protagonists in clear cases”. Sure, there is a statistically significant difference between knowledge and Gettier cases, but both a charitable interpretation of the quote and accurate statement of the present findings is that there is not much of a theoretically significant difference. As the authors acknowledge in their revisions, the simple truth of what the authors found is this: something like 82% to 88% (depending on order) of people attribute Gettier knowledge in Darrell cases as they would in paradigmatic knowledge controls (and approximately 31% in ignorance controls as they would paradigmatic knowledge). At the end of the day, there is a substantial amount of knowledge being attributed for salient failed threats does not look anything like ignorance. It is verifiably false that “the present research supports the view that nonphilosophers generally demonstrate Gettier intuitions”.

Our response: We acknowledge that the effect was small and may not be judged

as theoretically meaningful. We have added a sentence that makes this point on page 67 and on. We have also adjusted our language in a few places to decrease the strength of our conclusions.

We further agree that the statement that “nonphilosophers generally demonstrate Gettier intuitions” is not accurate and apologize for the misrepresentation of our results. We have changed our phrasing to state that “at least some non-philosophers demonstrate Gettier intuitions” (see page 70). Please note that we already explain in this paragraph that “Gettier intuitions were by no means common”, highlighting that a sizable proportion of participants denied knowledge in the knowledge condition and attributed knowledge in the Gettier condition.

Reviewer 4: “The remaining concern: Yes, Turri et al. suggested the Knowledge and Gettier conditions didn't differ. But they also noted that in both conditions, people mostly attributed knowledge. e.g., "When the threat failed to prevent Darrel from detecting the truth, participants attributed knowledge at rates exceeding chance..." (p. 381 in Turri et al.). The current MS also suggests that the rate of knowledge vs ignorance attribution should matter. For instance, the paper introduces Gettier cases as situations where "a person has a belief that is both true and well-supported by evidence... ...yet that person is not judged as possessing knowledge." (Also see the Abstract for more like this). But the issue of whether people overall attribute knowledge vs ignorance is backgrounded when discussing whether Turri et al was replicated, and in drawing the major conclusions. So, what was found? As in Turri et al. most participants (60%) attributed knowledge in the Darrell case (i.e., the case Turri et al used). In the Gerald case, knowledge attribution was 50%, which still doesn't really count as denying knowledge (as might be required for a Gettier intuition). And yes, people mostly denied knowledge in the Emma case, but they also did that in the Knowledge version of that vignette, so I don't think this means much. Contrary to what the authors claim, the paper showing that people don't robustly deny knowledge in lots of Gettier cases--and I believe this was a major claim of Turri et al.”

Our response: We agree that our research did not demonstrate a robust denial of knowledge in the Gettier cases examined. As explained in our response above, we have tempered our claims in the manuscript to better reflect the results.

1. Reviewer 4: The additions to the Introduction improve the paper, but my sense is that most readers will be confused. For example, I don't know see how anyone is supposed to know what "replacement-by-backup" or "replaced evidence" are supposed to mean from the added discussion on p. 10. The term "counterfeit object" is also unclear because dissimilar kinds of Gettier cases involve counterfeit objects:

-In Fake Barn type cases (i.e., the kinds of cases featured in this MS), the protagonist sees a real object but their knowledge is "threatened" because they could easily looked at a counterfeit object instead.

-Other Gettier cases use counterfeits in the opposite way: The protagonist initially sees a counterfeit item and forms a belief based on it. But that object is subsequently replaced by a real item, making the belief true by luck.

Both cases involve counterfeit objects, but the paper uses the term "counterfeit object case" to refer only to the first sort. So this is confusing. These issues again make me question the knowledge gain offered here. i.e., given the variety of Gettier cases out there, what's so important about focusing on this one kind of case?

Our response: We agree that the definitions of these terms could be more clear. We edited the paragraph on page 10 that begins “However, people do not demonstrate Gettier intuitions for some Gettier-type cases…” to include brief definitions of “replacement-by-backup” and “counterfeit-object” cases. We also clarified in a footnote that we are using the term counterfeit-object case in line with Powell et al. (2015) to describe what some researchers call “fake-barn” type cases.

We used this type of Gettier case because the report was a registered replication of Turri et al. (2015, Experiment 1), which used a counterfeit-object case (Darrel/squirrels).