

Final edits

2024-03-25

Dear reviewers and editors,

Thank you all once again for your very helpful feedback. I have compiled a single document this round in which I reply to your final suggestions and note where changes have been made. It has been a great experience to exchange these ideas, and I am very grateful to you all.

Dr. Plonsky

I have just one small but important point to add to the Handling Editor and the reviewer's comments:

I am with you 100% on, and have written on several occasions about, the misinterpretations of a lack of statistical significance. However, The leading part of the title is misleading in my view and should be changed, I think, UNLESS the main point of the paper is to argue against this fallacy of NHST, which I don't believe to be the case. As I see it, the main focus of the study is substantive in nature and the title should reflect this. The current title is not particularly informative in terms of telling potential readers what the study is about. Please consider re-wording to a more descriptive title.

Thank you for this point. I agree entirely. I have changed the title to "Statistical Insignificance is not wholesale transfer in L3 acquisition: An approximate Replication of Rothman (2011)" in hopes that this is more appropriate and informative regarding the content of the paper.

Dr. McManus

- Add to the abstract the main changes made to the initial study for the replication.

I reworked the abstract on the basis of this suggestion. I added a brief summary of the original study and what the replication changed (the materials, but not the population, question or analysis).

- Provide a definition for the replication approach (i.e., what is a conceptual replication study)?

I added this and after revisiting some of your work on replication research, I changed from conceptual to approximate. I argue this since the recreated materials are in the same format as the original (based on the example in the paper), and the same number of tokens per condition were used. In addition, the research question, population and analysis were all replicable.

- I noticed that the initial study's research questions were not used in the replication study. I recommend using the initial study's research questions and/or hypotheses, with appropriate and clearly identified edits that reflect any modifications made. If the new ones are retained, the text should state that these were not the same as those used in the initial study, with justifications provided.

The broad question is the same: how do the L1 and L2 impact L3 judgments? The new sub-question I ask if whether the conclusions of the original study will be replicated if we apply a post-hoc test of equivalence after also getting a null effect. I added clarity where I had previously left this implicit.

- State whether the initial study's data collection materials were available and used in the replication and, if used, how these were accessed. Procedures for determining the comparability of the new data collection materials in light of those described in the initial study should be outlined.

I added in a few places that these were not available.

Reviewer 1

I have some trouble understanding why the authors think it is important to specify how close is close enough or how different is different enough in the context of TPM. The critical point of the TPM is that transfer is from one language but not the other. Although perhaps a more comprehensive theory would need to consider to what degree transfer can happen, TPM's prediction is categorical. If that is the central point the authors wish to make, then it should be made clear and separately from arguing against Rothman (2011)'s results and interpretation purely from a statistical point of view. In any case, the results of Rothman (2011) provided no evidence against the TPM. If the concern is purely statistical, I encourage the author to make it clearer that the critical issue is not statistical procedure (being a specific theoretical perspective and/or power). Instead, researchers should interpret statistical outcomes with caution and accuracy.

Thank you for this comment. My aim is to argue that if we do not specify how close is close enough, that in addition to concerns with statistical power, we might not be able to provide clear support for either the TPM or LPM.

It is possible with increased sample size, for instance, to have a significance NHST and a significant TOST. If this were the case, it would be unclear which view would be supported without also specifying where meaningful differences begin.

Some minor comments: 1. All figures are outside the page margin.

Fixed.

2. The in-text citations do not follow APA 7 (as requested by the journal), e.g., For 3 or more authors, use the first author and "et al." for all in-text citations, etc.

Fixed via an APA7 csl file in the YAML frontmatter. The manuscript is generated in papaja which defaults to APA6.

3. Page 10, line 220 $p < .005$) => $p < .005$)
4. Page 10, line 223 $p < .005$) => $p < .005$)
5. Page 25, line 479 Kruschke (2018)) => Kruschke, 2018)
6. Page 26, line 513 Cabrelli and Iverson (2023)) => Cabrelli & Iverson, 2023)

Fixed all of these, thanks again!

Reviewer 2

Thank you for addressing in a way or another all my comments. I think this is a very important addition to the field and I am very much looking forward to seeing it published. I would only like to ask a bit more clarity when talking about CITH (page 25, line 507). The interpretation of the results and argumentations in relation to this is not easily understandable. If this paragraph could be reworked, it would be great.

Thank you for this feedback. I have reworked this paragraph to hopefully make more sense. Thank you!