Dr. Ljerka Ostojić

Recommender

PCI Registered Reports

January 17, 2022

Dear Dr. Ljerka Ostojić,

We are grateful for your honest handling and constructive comments on our manuscript.
Regarding your comment about a dependent variable "the filling rate of duration" and the experimental design, we have made amendments on our survey form to make it clearer. Moreover, as this point was described in detail in the response to Reviewer 2, please refer to that as well.

In addition, we have also corrected one point at the attention check question (ACQ). Originally, this study was supposed to be done by 2021 but now it turns out to be 2022. As a matter of fact, we changed the correct answer for ACQ to Reiwa 4 (2022). Thank you for your understanding.

Comments & Replies to Review by Chris Chambers

1-1
The authors have responded very thoroughly and impressively to my first review. The design is now much clearer to me (and in my opinion stronger), and the manuscript is just about ready for Stage 1 IPA.
I have only one further question, related to the revisions that have now been made. For H2-1 and H2-2, the authors propose analysing only a sub-sample of the total sample size based on a power analysis suggesting that N=102 and N=386 would be sufficient, respectively. This power analysis is appears to be based on a generic d (or dz) of 0.4, as advocated by Brysbaert (2019 -- note that this reference is not in the reference list, so I am not sure which article exactly it is referring to; on an additional minor note, please also ensure that you use "dz" for the within-subjects effect sizes and "d" for the between-subjects effect sizes).
Possibly I am missing a key point here, but my main question is: given the generic nature of the 0.4 effect size estimation, why not take advantage of the total sample size in each study to also test H2-1 and H2-2? It is true that H1 requires a much larger sample size than H2-1 and H2-2, but regardless it seems a shame to leave all of that additional data unanalysed when it would be diagnostic about the predictiosn, and when including this extra data will simply have the benefit of making the statistical tests for H2-1 and H2-2 more sensitive to effects smaller than 0.4 (which I assume would still be theoretically relevant). The authors could report an a priori sensitivity power analysis reporting what effect size they have 0.95 to detect for H2-1 and H2-2 given the full sample size (presumably much smaller than 0.4).

Reply:
Thank you for your suggestions. We are so sorry that we missed an important reference and now it has been added to the reference list. Also thank you for pointing out the "*d*" in our within-subject design, we realized that it should be distinguished from *d* to prevent misunderstanding and have already changed it from "*d*" to "*dz*". We conducted an a priori sensitivity power analysis reporting that *dz* = 0.13 in Study 1 (*dz* = 0.11 in Study 2) for H2-1 and critical *r* = .076 in Study 1(*r* = .064 in Study 2) for

H2-2. These results are apparently smaller than 0.4 as we are grateful to accept your suggestion to take advantage of the total sample size in each study.

1-2
My only other comment is that in the study design table, the content of the columns "Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis" and "Interpretation given different outcomes" should be adjusted slightly to fit requirements. The column "Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis" should provide a justification of the smallest effect size of interest and power level for each hypothesis test, rather than a description of what significant differences would indicate (as currently). The existing content for this column should instead be combined with the existing content in the "Interpretation given different outcomes" column and then included solely within the "Interpretation given different outcomes" column.

Reply:
We are grateful to receive your thoughtful suggestions about our design table. We have adjusted the content of these two columns to fit requirements.

1-3
Minor: in the third column of the study design table there is an inconsistency in line spacing.

Reply:
Thank you for your detailed check and we have modified the line spacing to be consistent with other columns.

Comments & Replies to Review by Anonymous Reviewer

The study seems well designed to test the intriguing and plausible alternative hypothesis raised by the authors. The hypotheses and methods - ostensibly the most important aspects of a registered report - are very clear and reasonable. I only have a few queries and comments (point 3 being the most important):

Reply:
Thank you for your kind and detailed comments. We have responded to each of your comments as follows.

1. The authors identify previous attempts to explain the TDE from perspectives other than spatial movement (Gan et al., 2017; Mrkva et al., 2018; McCormack et al., 2019), but don't describe these alternative explanations in any detail. It would be helpful to further clarify these explanations in order to illuminate the novel contribution of the authors' own hypothesis.

Reply:
Thank you for pointing this out. We have described in more detail the hypotheses/interpretations raised in previous studies to make the contributions of the present study more clearly known to readers.

2. The authors write that the past comprises both "predetermined and sudden events", but it would seem that "expected and unexpected events" would be more appropriate in this context.

Reply:

Thank you for confirming not only the content, but also the expression. As non-native English speakers, we appreciate it very much. We corrected it as you suggested.

3. **The authors claim that, unlike in previous studies, they will focus on the length of intervening events as well as the absolute number of events. Yet, their Likert scales don't seem to capture event length in any precise manner. How is this factor being incorporated into the current study, and how does this differ from previous studies?**
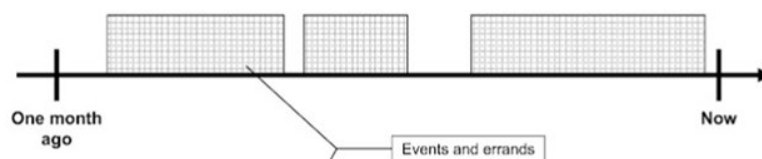
Reply:

Thank you for pointing it out and in order to make a clearer explanation for our assumption, we introduced figures to make it easier for participants' understanding as follows. Please also refer to Supplementary Information for the full instruction.

How much has your schedule been filled with errands and events for the last month? *
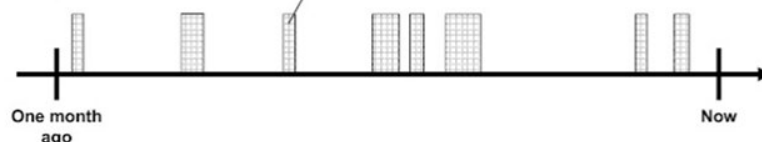
Examples of durations filled by events and errands
(As shown in the examples below, please consider both the number and length of each event/errand when answering.)

An example of the filled duration

One month ago — Now

Events and errands

An example of the unfilled duration

One month ago — Now

1   2   3   4   5   6   7   8   9   10

Not filled at all.   ○ ○ ○ ○ ○ ○ ○ ○ ○ ○   All filled up.

4. It is unclear why the authors would limit their analyses to a small fraction of the sample when testing some of their hypotheses. Why not include the full sample to get a better estimate of the true effect? Perhaps there is a good reason for such a statistical practice, but if so the authors should clarify it.

Reply:

Thank you for your suggestion. Regarding your suggestion, we have changed our plan to analyze the full sample. We will then conduct sensitivity analysis and report the results.

5. The sample will include residents of Japan, whereas the Caruso et al. (2013) studies seems to have included American undergraduates and M-Turk participants. Are there any potential cross-

cultural differences in time perspectives that could produce differences in the findings, independent of the authors' hypothesis? Perhaps this is simply an unavoidable limitation that would need to be addressed in the Discussion.

Reply:
We are aware that cultural and linguistic differences need to be considered. Thank you for the hint to enrich our discussion.