

Dear Prof. Chambers,

Thank you again for gathering another round of reviews in such a short time. We have addressed all the reviewers' remaining points below. Please, let us know if we can move forward with our manuscript – We are looking forward to starting data collection for pilot 2.

Many thanks,

Eva (on behalf of all the authors)

Reviews

Review by Natalia Kartushina, 18 Mar 2022 14:26

- 1) I would like to thank the authors for providing such a thorough and clear rebuttal letter. They have satisfactorily addressed all my comments and questions and those of my colleagues, I believe. I was pleased to see a schematic illustration of the stimuli configuration and to have an opportunity to try out a new version of the game. I would like to mention that it is an honor and a privilege to be able to contribute to this absolutely high-standard piece of work! I have only one remaining concern that I wanted to bring to the authors' attention in an attempt to reflect on it, but not necessarily to require any changes.*

Thank you for your kind words, we are glad that Dr Kartushina found our responses sufficient to solve most of her concerns.

- 2) I appreciated that the authors explained why they added the distractor objects in the design; yet, it seems to me that children in the LV group would quickly realize the lack of variability in stimuli (and what are the target objects) and would ignore the distractors, which might not (fully) serve the purpose of having them, i.e., mitigating against associating images with specific moves without listening to the sentences for the LV group. For the HV group on the other hand, having the distractor objects can make the task more difficult and decrease their chance level, as they might consider the distractor objects to a larger extent as compared to the LV group. This brings us back to the original concern with respect to the chance level and the trial removal strategy (see below). This does not have major implications for the analyses, but maybe this needs to be thought thoroughly before the data collection, as it can affect the chance level, which is lower than 25% then. Maybe the authors could comment on this?*

The chance levels are computed in terms of the number of possible responses for the dataset under consideration. It is true that if children do not consider some options, then in a sense their "chance" level is different, but this is essentially the same as saying performance is lower. As noted previously, the fact that the LV group will more quickly focus on certain cues than the HV group is part of the difference between the learning conditions.

- 3) I am afraid I have doubts about the exclusion of the distractors-related moves from the analyses. The authors indicate that these moves mean that children haven't identified the two nouns involved in the sentence; however, given that the pilot data demonstrates that children learnt the words at the noun practice task and the authors will exclude participants who would have scored below 80% in the noun practice task, then we would expect only few of them to*

occur (although see the above concern), so why not keeping these moves in the analyses (as they also contribute to decreasing the chance level)?

We have had a lot of discussion as to whether to exclude these or not. One advantage of excluding them is that it makes the analyses more directly comparable to those we ran for the pilot 1 data, and thus makes the priors and sample size computations from that data more relevant. We do think it is possible there might be interesting data in the “excluded trials” and for that reason we have also proposed the additional series of analyses in the “follow-up analysis” section which we will run alongside our main analyses.

- 4) *Relatedly, the authors indicate that they will measure verbal short-term memory and will include this in the analyses “as a covariate which is expected to explain variance in learning over the training sessions”; wouldn’t then it benefit the analyses to not exclude from the model trials where children made errors (i.e., with the distractors), as they are likely related to differences in verbal short-memory (among others)?*

Our intention is not to include verbal-short term memory as a covariate in every analysis, but instead only in exploratory analyses (see page 33, lines 906-911). Following your point (5) below, we think it would make sense to do this *both* for the main core analyses *and* for the series of follow-up analyses below. We agree that seems likely that these relationships may turn out to be more important in the analyses looking at distractors.

- 5) *The inclusion of the analysis “Recognition of nouns in continuous speech—Do children correctly identify the nouns in the sentence?” is a very nice way to address (potential) differences in the proportion of distractor-related errors between conditions. I wonder whether the verbal short-term memory shall be included in the model as a co-variate.*

See response to (4) above.

- 6) *In this analysis, the authors set the chance level to 1/18; yet, I wonder whether this shall be changed (increased) given that four (A,B,C,D) out of 18 moves are counted as accurate.*

Yes – thank you for spotting this. Chance should indeed be 4/18 and this was an error in the previous manuscript. We have now corrected this (page 38, lines 1014).

- 7) *Minor comments/questions:*

a. *Line 17: remove “an” (abstract relationships)*

Thank you, this has been already flagged as removed in the current version of the manuscript (i.e., “n”). In the final version track changes will be turned off so that additions and deletions will be easier to see.

b. *Line 996: set movable => set of movable*

c. *Figure 6 caption: moving the banana to square 6 => moving the banana to square 4*

Thank you, these have been corrected now.

d. *Why the memory span task was run twice?*

We plan to run the working memory task in the proposed experiment twice, before and after the experiment, following the approach taken by Parker et al., 2022, which is the RR replicating the original Hsu & Bishop, 2014, and from which we have taken the working memory task. The goal is to obtain a more robust measure of working memory. We have added this information to the manuscript (p 41, line 1068-1069).

e. *Line 1275: the moves B, C, D are repeated twice*

These are repeated twice but in two different sentences:

1st sentence: "We consider for the main analyses [...] referred as A, B, C, D. "
2nd sentence: "A-moves are accurate, while B, C, D are inaccurate."

f. *Line 1844: "the majority of children (i.e., >5)", do you mean more than 5 children would be majority?*

Since the pilot aims at collecting data from 10 children, with the "majority of them" we mean more than 5.

8) *Best of luck with Pilot 2! and again congratulations on such an exemplary registered report!*

Thank you very much, hopefully we will hear back from you in stage 2.

Review by Caroline Rowland, 24 Mar 2022 07:04

9) *I want to thank the authors for paying such close attention to my comments; this makes the process of reviewing feel worthwhile! The authors have responded well I have no hesitation in recommending acceptance. I agree that it will be interesting to include the working memory task. I also thank them for providing a more detailed explanation of how moves will be removed from analysis and for the interesting new analyses on word order and on differences earlier in learning. I am looking forward to seeing the results.*

Thank you for your kind words, we are happy that Prof. Rowland has found our response to her points adequate.

10) *The comments below are merely for consideration when writing the stage 2 report, after data are collected and analysed. Regarding noise and the choice of priors, the authors have provided sound justification for their decision. If they do find null results, however, I would urge them to discuss the choice of prior as a possible reason for this in the discussion and, if possible, dismiss it (by for example, providing the same type of comparison of SDs that they provide in their reply)*

Thank you. We intend to include robustness regions which will let us see the range of priors which would have led to the same conclusions, and we will indeed consider this in the discussion.

11) *Footnote page 10: "From a theoretical perspective, although transfer across related constructions does happen in some cases (Abbot-Smith & Behrens 2001) we do not believe there is any good reason to expect strong transfer to new adpositions in this paradigm". I still didn't quite follow why - from a theoretical perspective - the authors do not expect strong transfer, as a clear explanation wasn't given. But this could be something to address further in the discussion - under what circumstances the authors would, and would not, expect transfer.*

Thank you for the suggestion which we agree is best addressed in discussion.

Review by Julien Mayor, 10 Mar 2022 09:02

The authors have been extremely responsive in addressing comments raised in the first round of reviews.

I only have two minor comments:

12) *The new design is very nice, in addressing the issue of word order. I just wonder if this will be clear, for participants, that they should take an object from bottom of the screen and move it into the grid (apologies if I have missed it)? As the object to be moved is also present in the grid itself, my first attempt was to move that object instead. I guess that, if participants are providing an erroneous move, they will be illustrated with the correct move, and hence they will ultimately understand what they are required to do - but that would mean that during the first move (or couple of moves) participants may still be learning the rules of the game, rather than about cracking the sentences' meanings (and if that's so, a few practice trials would be great).*

We hope that this won't be a problem given that, unlike in the link we provided, prior to starting, the game children will watch a brief demo video which shows them how to move the object on the grid and gives some examples (see page 42 line 1087-1091). Our pilot 2 data will further establish that children are able to play the game.

13) *On line 939 "with respect to which the first noun should be moved" should probably read "with respect to which the first object should be moved"*

Thank you, this has been corrected in the new version of the manuscript.

