

## **Peer review comments**

**Article title:** Moving developmental research online: comparing in-lab and web-based studies of model-based reinforcement learning

**Author names:** Kate Nussenbaum, Maximilian Scheuplein, Camille V. Phaneuef, Michael D. Evans, Catherine A. Hartley

**Handling editor:** Kate Mills

## **Editor first decision – Revise & Resubmit**

August 27, 2020

Dear Kate Nussenbaum,

I have now received all reviews of your manuscript, “Moving developmental research online: comparing in-lab and web-based studies of model-based reinforcement learning” from qualified researchers, as well as read the manuscript myself. As you can see in the reviewers’ comments, there is overwhelming enthusiasm for this study and manuscript. I am also in agreement that this study is sound and manuscript clearly written, describing an exciting contribution to developmental science! I must admit I am slightly biased as a parent of a child who has participated in studies advertised through ChildrenHelpingScience (though not the present study described in the manuscript), and found this paper fascinating from both a scientific perspective but also as a parent-in-the-background. The reviewers have also provided constructive feedback, which I agree will strengthen the manuscript. I therefore encourage you to submit a revised version for further consideration at Collabra: Psychology.

The reviewers did an outstanding job in their reviews. I include the reviewers main suggestions here in this message, but I encourage you to read the full reviews in the Scholastica portal. In your resubmission, please include a document with a point-by-point response to both the points I list here and any additional reviewers’ comments, outlining each change made in your manuscript or providing a suitable rebuttal.

Reviewer 1: p. 4: The description of Decker et al (2016), and the brief discussion of the prior in-lab replication of it, would benefit from a simpler and clearer statement of what the task is and what general questions in the sub-field the task is used for. The task is described in full detail in the Methods, but a quick summary here would make the introduction a bit more accessible to developmental scientists unfamiliar with decision-making research. This particularly because the paper will likely be cited as evidence in favor of web-based developmental research in other domains.

p. 6: In my lab’s studies of children on the web, we often just ask parents if they did interfere at any point; the parents sometimes volunteer that they did interfere for one reason or another. I wonder if an item like that was used (or maybe should be recommended)?

p. 8: It's cool that some families were recruited from Children Helping Science, which, as far as I know, hasn't had data published about it (yet). Are there enough families recruited from there that the subset of CHS families could be compared to non-CHS families, as a test of the efficacy of that recruitment method?

fig. 1: I think this figure could be a bit clearer. In Panel A, It would help to make a more explicit separation between the present results "Adults" and "Dev" from what they're being compared to, so that the comparison is more obvious; they should also be split into "Adult", "Adolescents", and "Children" to match the other figures. Then, in panels B and C, why not use the same style of stacked bar graph, rather than the histograms, so that the reader doesn't need to learn a second type of visual representation?

fig. 2: This information is in the text, but it might be helpful to include a schematic in Fig 2 depicting the progression of the study (X happened for N trials, then Y happened, etc).

p. 16: I will probably cite this paper (and I think others will too) as an example of the relatively low rates of "bad behavior" in online studies. As such I think it would be a good idea to report tables with the actual rates of "bad behavior", in addition to the histograms in figs 4 and 7 (which are helpfully illustrative, but imprecise for this purpose). That information could go in SI, or maybe be added to Table 2.

pp. 17-18: One analysis question: were age effects tested with age as a continuous variable, or as three categories? "We also observed a main effect of age ( $p_s < .001$ )" is ambiguous, and visualizations suggesting either type of analysis are presented in Fig 5. I have a preference for using age as a continuous variable, but it's also reasonable to repeat the categorical comparisons done in previous studies (and helpful to visualize those as currently visualized). Please clarify.

Fig. 9: The six atypical observations (below WASI of ~15) look like they might be influential observations, skewing the interpretation of the relation between WASI and MaRs-IB. Seeing as this analysis is fully exploratory anyway, I think it might be worth running a sensitivity analysis that excludes or Winsorizes them and measures the impact on the regression results, just to be sure that the reported relation holds. It looks like it might actually be underestimated because of those observations, but I'm not sure.

Reviewer 2: First, I think that the authors should make their data publicly available online if at all possible.

Second, I think Figure 1 is hard to read. Not only is the font size in this figure too small, I also the ordering of the bars makes it hard to compare the online and "regular" studies. I think it would be useful to plot the adult distributions of both the current and Potter studies side by side, and do the same for the developmental samples. Also, I think that it

would be useful to use the label ‘online’ for the bars corresponding to the samples in the current study (mirroring the use of this label in subsequent figures).

Third, I think the paper would benefit from a figure that illustrates the transition structure of this task, as well as depictions of the stay probability bar graphs for purely model-free and model-based agents. I think such figures would be beneficial for readers who are interested in doing developmental research online, but who are not familiar with this reinforcement learning task.

---

In summary, I think this is a promising manuscript and, I hope you will revise it for further consideration at Collabra: Psychology. I look forward to receiving your revision!

Please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained. This is the last opportunity for major editing, therefore please fully check your file prior to re-submission.

If you have any questions or difficulties during this process, please contact the editorial office at [editorialoffice@collabra.org](mailto:editorialoffice@collabra.org).

We hope you can submit your revision within the next six weeks. If you cannot make this deadline, please let us know as early as possible.

Sincerely,

Kate Mills

## Reviewer 1

### Open response questions

Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

Review of “Moving developmental research online: Comparing in-lab and web-based studies of model-based reinforcement learning”

This manuscript reports a replication of a prior decision-making result (run in-lab) in a sample of children, adolescents, and adults (run online) and analyses of the similarities

and differences in results across the recruitment methods. The results provide convincing evidence that web-based experiments in decision-making research are externally valid. The bigger implication, which is hinted at in the last paragraph of the manuscript, is that developmental science in general is amenable to web-based research. I enjoyed the paper and, while I am hardly unbiased (I do web-based research myself, and help others do it) I didn't see anything in the manuscript that I took major issue with.

Below are a few suggestions which I think would help round out the paper in a revision.

p. 4: The description of Decker et al (2016), and the brief discussion of the prior in-lab replication of it, would benefit from a simpler and clearer statement of what the task is and what general questions in the sub-field the task is used for. The task is described in full detail in the Methods, but a quick summary here would make the introduction a bit more accessible to developmental scientists unfamiliar with decision-making research. This particularly because the paper will likely be cited as evidence in favor of web-based developmental research in other domains.

p. 6: In my lab's studies of children on the web, we often just ask parents if they did interfere at any point; the parents sometimes volunteer that they did interfere for one reason or another. I wonder if an item like that was used (or maybe should be recommended)?

p. 8: It's cool that some families were recruited from Children Helping Science, which, as far as I know, hasn't had data published about it (yet). Are there enough families recruited from there that the subset of CHS families could be compared to non-CHS families, as a test of the efficacy of that recruitment method?

fig. 1: I think this figure could be a bit clearer. In Panel A, It would help to make a more explicit separation between the present results "Adults" and "Dev" from what they're being compared to, so that the comparison is more obvious; they should also be split into "Adult", "Adolescents", and "Children" to match the other figures. Then, in panels B and C, why not use the same style of stacked bar graph, rather than the histograms, so that the reader doesn't need to learn a second type of visual representation?

fig. 2: This information is in the text, but it might be helpful to include a schematic in Fig 2 depicting the progression of the study (X happened for N trials, then Y happened, etc).

p. 16: I will probably cite this paper (and I think others will too) as an example of the relatively low rates of "bad behavior" in online studies. As such I think it would be a good idea to report tables with the actual rates of "bad behavior", in addition to the histograms in figs 4 and 7 (which are helpfully illustrative, but imprecise for this purpose). That information could go in SI, or maybe be added to Table 2.

pp. 17-18: One analysis question: were age effects tested with age as a continuous variable, or as three categories? “We also observed a main effect of age ( $p_s < .001$ )” is ambiguous, and visualizations suggesting either type of analysis are presented in Fig 5. I have a preference for using age as a continuous variable, but it’s also reasonable to repeat the categorical comparisons done in previous studies (and helpful to visualize those as currently visualized). Please clarify.

Fig. 9: The six atypical observations (below WASI of ~15) look like they might be influential observations, skewing the interpretation of the relation between WASI and MaRs-IB. Seeing as this analysis is fully exploratory anyway, I think it might be worth running a sensitivity analysis that excludes or Winsorizes them and measures the impact on the regression results, just to be sure that the reported relation holds. It looks like it might actually be underestimated because of those observations, but I’m not sure.

Thanks to the authors for their hard work on an interesting and useful paper.

-Samuel Mehr

## Reviewer 2

### Open response questions

Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.

The authors report a study that aims to investigate whether it is feasible to collect online data for experiments in developmental psychology. As a case study, they attempt to replicate the results of Decker et al. and Potter et al., who showed that (a) the degree of model-based decision making in the two-step task increases from childhood into adolescence, and (b) that this effect is mediated by abstract reasoning ability in an online sample.

The results are straightforward – the authors seem to replicate all these effects. Moreover, throughout the paper they provide valuable insights into how to set up an online environment conducive for running psychological experiments in developmental samples.

I have never made this recommendation, but I think this paper should be published virtually as-is. The paper is well written, the analyses are thorough and appropriate, and the results are important. I also think that description of all the precautions the authors took to facilitate online data collection in developmental studies will be invaluable for

other developmental labs that are interested in (but not experienced with) online data collection.

That being said, I have three recommendations.

First, I think that the authors should make their data publicly available online if at all possible.

Second, I think Figure 1 is hard to read. Not only is the font size in this figure too small, I also the ordering of the bars makes it hard to compare the online and “regular” studies. I think it would be useful to plot the adult distributions of both the current and Potter studies side by side, and do the same for the developmental samples. Also, I think that it would be useful to use the label ‘online’ for the bars corresponding to the samples in the current study (mirroring the use of this label in subsequent figures).

Third, I think the paper would benefit from a figure that illustrates the transition structure of this task, as well as depictions of the stay probability bar graphs for purely model-free and model-based agents. I think such figures would be beneficial for readers who are interested in doing developmental research online, but who are not familiar with this reinforcement learning task.

Typo:

In the reference section, the Feher Da Silva & Hare paper has some formatting errors.

### **Authors response letter**

August 31, 2020

Dear Dr. Mills,

We are submitting a revised version of our manuscript, “Moving developmental research online:

comparing in-lab and web-based studies of model-based reinforcement learning” for consideration as an empirical article in *Collabra: Psychology*.

We very much appreciate your and the reviewers’ time and thoughtful comments. All comments are addressed in turn below. We have incorporated Reviewer 1’s suggestion to expand the

description of the two-step task in the introduction to make the paper clearer to a broader audience. In service of this same goal, we have also incorporated Reviewer 2’s suggestion to provide a figure depicting the task’s transition structure and the expected choice patterns for pure model-free and model-based learners.

Again, we appreciate your work in helping us improve this manuscript, and we look forward to

hearing from you.

Sincerely,

Kate Nussenbaum

Maximilian Scheuplein

Camille V. Phaneuef

Michael D. Evans

Catherine A. Hartley

See a separate file in the supplemental material directory.

**Editor final decision**

August 31, 2020

Dear Kate Nussenbaum,

I have now had a chance to read over your manuscript “Moving developmental research online: comparing in-lab and web-based studies of model-based reinforcement learning”, along with the letter describing the changes you made. Thank you for your responsiveness to the concerns that the reviewers and I raised. I am happy to say that your paper is now officially accepted for publication in Collabra: Psychology. Congratulations on this excellent work, I think it will make an important contribution to the literature and I look forward to seeing it published! I hope your experiences with Collabra: Psychology have been positive and that you will continue to consider it as an outlet for your work.

As there are no further reviewer revisions to make, you do not have to complete any tasks at this point. Our managing editor will contact you in case there are any pre-production file related questions. You will have an opportunity to check the page proofs before we publish your article. Thank you again for publishing in Collabra: Psychology.

Sincerely, Kate Mills