Response to reviewers of JCGS-23-236

Perception and Cognitive Implications of Logarithmic Scales for Exponentially Increasing Data: Perceptual Sensitivity Tested with Statistical Lineups

> To send to Susan during meetings: https://earobinson95.github.io/logarithmic-lineups/revision-responses-1.pdf
>
> Revised manuscript: https://earobinson95.github.io/logarithmic-lineups/logarithmic-lineups-revisions.pdf
>
> For Emily to submit: https://mc.manuscriptcentral.com/jcgs?URL_MASK=f1cfe0d4da184e54b91df29f43be0038
>
> Your paper will be printed in black and white. The online version will be in color but to accommodate printing without color please be sure you don't refer to color in either the figure captions or text and that all figures are still interpretable in black and white.

All responses to reviewers' comments are written in blue. A `diffs-1.pdf` file has also been included to show changes.

Need to address yet. Double check / review.

Reviewer: 1

Summary: This manuscript describes a study that assessed peoples' abilities to distinguish data with different curvatures. Using statistical lineups, the study found that it was similarly easy to pick out differences in curvature on linear and log scales when the differences were large, but when they were small it was easier to pick out differences on log scales when the target plot was less curved than the null plots.

Significance and novelty: I'm not an expert in cognitive psychology, so I don't know that literature well, but based on my limited understanding, this study is novel, and I find it interesting.

Quality of work: From what I can tell the work is well done, and with some revision I think it will be really solid. In particular, I think including results from the Rorschach lineups and including results on the confidence and justification questions would make this substantially more rigorous. See suggestions for improvement below for more detailed feedback on how it could be improved.

Fit: This work fits well within the scope of JCGS.

Suggestions for improvement:

There are three categories of suggestions, ordered here from issues that would take the most time and effort to those that would take the least.

The first category is which information to include in the manuscript.

- The manuscript and the applet make it clear that there are two other parts of this overall study. For a number of reasons, I think it would improve the work to include all three together in a single study rather than splitting them into three separate papers. The strongest reason is that it gives a more complete picture. The writing of the present manuscript makes it clear that log vs. linear presentation of data is a main focus, but the results from this first study speak more to the degree of curviness observed on a panel than to anything specifically about log vs. linear scales per se. It seems like the second and third pieces will focus more on log vs. linear scales, in which case you could make a more compelling case with all of them together. Another reason is that the current study is not particularly long, so I don't think it would be overly taxing to readers to including one or two more sets of results. I understand that there are reasons to split them up as well, so I'm happy to leave this decision to the authors (and to some degree to the editor), but I wanted to mention it in case you agree.

We appreciate your review of our current manuscript and recognition that this is a part of a larger study. While we agree that the series of studies shows a comprehensive perspective on the perception and use of logarithmic scales, it is important to note that each study has undergone its own unique development process and analysis of results. Combining all of these studies into a single manuscript would significantly increase the length of the paper, a practice not reccommended by JCGS editors during the revision process. Therefore, we have decided to continue with our original plan, which involves publishing the three studies individiually in JCGS over the course of the next year. Ultimately, we will conclude this series with a final paper that references and synthesizes the findings from the preceding studies. THis approach aligns with common practices in fields such as psychology.

- For the study as written, I suggest including your data from the Rorschach lineups. You have the data, and they're good controls for the overall study.
- The methods indicate that participants were asked not only about which panel differed, but also about their confidence and their justifications. I would find the confidence and

the justification data illuminating, and I suggest including results from those questions in addition to the right/wrong results. This applies to the Rorschach lineups as well.

- "In each lineup evaluation, participants justified their choice and provided their level of confidence in their choice." It would be good to see in the paper exactly how these questions were phrased and what the response options were, as these details matter for surveys.

Figure 7: Screenshot

- For presenting the results, I would find it very useful to see the % correct in the different combinations as well as the odds ratios. The overall % correct is given at the beginning of the results, but the breakdown for the different pairings would be even more informative.

https://www.tandfonline.com/cms/asset/45b880a8-c44d-4985-af25-a4de985a7d91/ucgs_a_1209116_f0010_b.gif

The second category is the interpretation of the results, which might require some editing, but doesn't require any major additions of results.

- The introduction focuses on the role of log vs. linear scales with respect to exponentially increasing data. As the text points out, this is relevant in a number of contexts. From what I can tell, though, the study is really testing for participants' abilities to see different degrees of curvature rather than anything specific to log vs. linear scales. (That is, there's nothing in the lineups that shows numbers on an axis or anything else that indicates the scale.) For instance, it would be possible to create curves that aren't exponential that have the same degrees and differences in curviness, and they would presumably give the same results. I'm not suggesting you do so; rather, I'm trying to point out that the results might be more broadly applicable to "curviness" than to log vs. linear scales alone. I think it's fine to motivate the work with the case of log vs. linear presentation of exponential data (particularly if you're combining with the other two parts of the overall study), but I suggest interpreting more broadly in the context of the degree of curvature displayed on a panel.

The third category is the details of presentation. In most cases the presentation is very clear, but there are a few places where some editing would improve the clarity.

- "In this paper, we evaluated the benefits and drawbacks of using log scales and examine their impact on perceptual sensitivity" The intro funnels pretty well, but the first sentence of the "In this paper" paragraph (which isn't the last) doesn't really clarify the specific goal. Perceptual sensitivity of what? It later becomes clear that the focus is on the degree of curvature, but it would be good to clarify that earlier in the paper (before this paragraph).
- Fig. 3 caption and analogous text: I'm confused about Fig. 3 in (I think) two separate ways.

  - The first way I'm confused is what's shown on the panels. The caption says "increasing exponential data" but doesn't give a formula (unless (2x5) + 3 or 2x2 is the formula, but that doesn't make sense to me). Given the description of what a lineup is, my understanding is that there should be two separate things shown in each side of the figure: 19 panels should show random data from the same formula, and the target should show something else. I'm not sure what the something else is, but it should be somehow different. Are the increasing exponential data the 19 panels or the 1? I suspect the 19, given that one panel (panel 4) looks decidedly non-power-law on the right-hand lineup, but it should be clear from the caption.
  - The second way I'm confused is what the "(2x5) + 3" and "2x2" represent. The text says they denote panels – does that just mean panel 13 (2x5 = 10 + 3 = 13) and 4 (2x2 = 4)? I suspect it might because panel 4 on the right is what looks different (as in, not exponential data, like the others appear to be). If so, though, why not say panel 13 and panel 4?

This is correct, in looking at prior literature presenting lineups, authors typically would include some mental math requiring participants to inspect which plot might be the target panel before immediately revealing the answer. We have adjusted our manuscript to reveal the panels as 13 and 4 per the reviewer's reccommendation.

- If these are examples from the actual versions shown to viewers, these make a lot more sense after reading the methods, but if the figure is presented where it is in the paper it needs a little more explanation. More clearly labeling which panel is the target and giving the formulas would help.
- Algorithm 1: I get the general idea, but the specifics are unclear. For example, point 1 says: "Determine the y = -x line scaled to fit the assigned domain and range" I'm not sure what this means. The y = -x line sounds like a line with intercept 0 and slope -1, but I'm not sure how that could be fit to the assigned domain and range since there aren't any free parameters.

This is correct, the y = -x line is a line with intecept 0 and slope -1. This was used to select the midpoint on the x-axis. We then jittered a small amount in both directions to have 4 points which are then used to obtain estimates for α, β, and θ. See https://srvanderplas.github.io/Perception-of-Log-Scales/lineups-development/simulations/model-options.html.

- Page 12: I don't know what the age of majority is.

We added clarification that the age of majority is 18+ in most regions and 19+ in certain U.S. states. This was the age in which participants could legally consent to participation in the study based on our IRB.

- Eq. 2: Does the Y on the left-hand side need subscript l as well as ijk?

Yes, thank you for this catch; the subscript l indicates the participant.

- Fig. 8 caption typo: the the

Thank you again, for the typo catch; we have updated the caption to include only one "the".

Reviewer: 2

This manuscript investigates perception of differences between linear and exponential functions on linear and log and graphs. I am not an expert in the field of perception, so ultimately, I am uncertain as to whether this should be published because I am not familiar with the level of novelty or importance of the results within that field. The study itself does seem to be robust and the results, though somewhat incremental, are clear and worth archiving. I do have some comments that could be addressed.

1. The characterization of the visual shape and functional behavior of an exponential function was not adequate, because it was not differentiated from functions with similar behaviors. For example the statement: "Early stages of exponential growth often appear to have a small growth rate, while the middle stage seems to exhibit more quadratic growth" describes many functions, including power law functions. So it seems that "general misinterpretation of exponential growth" is not clarified in this paper. It is not unreasonable to consider that many people quantitatively and qualitatively confuse exponential growth (e^x) with power law growth (x^n). A better overview of exponential growth and it's uniqueness compared to other common functions and importance in science/medicine seems warranted.
2. Expanding on #1, if one considers power laws, then log-log plots are commonly used and may apply to this area of study. That is one could compare linear and log-log plots for linear and power law functions. Maybe the overall results are different in this case compared to your study? That is, you speak of variation in curvature within the domain of exponential functions, but is the differing shape of power law curves important? I realize this is outside the scope of this paper, but acknowledging this boundary/limitation (and potential next step) of your study may be worth mentioning. Maybe not all high curvature curves behave the same perceptually? Also, extending this to power law functions could broaden the importance/usefulness of your results.
3. Using a VonBergmann as a reference is a little misleading in the sense that citations are usually used for empirical results of research.

Using comics that appeared contemporaneously with the events that occured with this study, provides a real world demonstration of the fundamental content we are exploring here. As such, they are a concise way of providing reasoning of the evidence we are exploring. And they clearly connect the perception of curvature and exponential growth.

4. Sections 2.1 and 2.2 were difficult to understand, very underspecified. It seemed to me that more explanation is needed about the algorithms and the parameter selection. For example, in algorithm #1, the y values are described as input parameters. How were they generated via eq 1? Eg, how were beta and sigma chosen? Maybe this is all standard for this field, so your description is adequate, but there were a few missing pieces for me that prevented me from understanding about the data sets were generated.
5. Students were asked about their level of confidence in their choice. How was this used? Maybe I missed it. It seems like it should be used/modeled.
6. How are these results unique, eg compared to Best et al.? IS this just a replication of results?
7. It would be good to talk about limitations. For example, to what extent is this only about perception comparisons of a singleton in a field of examples? How relevant is that to practice? For example, looking at only one graph (and determining whether the line is straight) or comparing only two graphs?
8. Tabulate the results (e.g., estimates, SEs, fit stats etc) of the GLMM
9. For fig 8, instead of odds ratio, why not give something a little more intuitive, like the probability of detection of target? Fig 8 is nice in that it holds a lot of info, but having probabilities is more concrete and useful.

Editor's Comments to Author:

Please pay particular attention to the comments from one reviewer about whether your work has implications for log-transformed data or for curves more generally. Also please provide a response to the request for more information about the survey questions and options and the data on confidence and justification.

SW - Major Revision