Perception of exponentially increasing data displayed on a log scale

by Emily A. Robinson, Reka Howard, Susan VanderPlas

December 7, 2020

1 Introduction

Graphics are a useful tool for displaying and communicating information. [Vanderplas et al., 2020] Researchers include graphics to communicate their results in scientific publications and news sources rely on graphics to convey news stories to the public. At the beginning of the SARS-NCOV-2 pandemic, we saw an influx of dashboards being developed to display case counts, transmission rates, and outbreak regions [Charlotte, 2020]; mass media routinely showed charts to provide information about COVID-19 [Romano et al., 2020]. People began seeking out graphical displays of coronavirus data as a direct result of these pieces of work [Rost, 2020]; providing increased and ongoing exposure to these graphics over time. Many of these graphics helped guide decision makers to implement policies such as shut-downs or mandated mask wearing, but also facilitated communication with the public to increase compliance [Bavel et al., 2020]. With the increasing importance graphics play in our everyday lives, we must actively choose which of many possible graphics to draw, according to some set of design choices to ensure that our charts are effective, as suggested in Unwin [2020].

When faced with data which spans several orders of magnitude, we must decide whether to show the data on its original scale (compressing the smaller magnitudes into relatively little area) or to transform the scale and alter the contextual appearance of the data. One common solution is to use a log scale transformation to display data over several orders of magnitude within one graph. Logarithms make multiplicative relationships additive, showing elasticities and other proportional changes, and also linearize power laws [Menge et al., 2018]. When presenting log-scaled data, it is possible to use either untransformed scale labels (for example, values of 1, 10 and 100 are equally spaced along the axis) or log-transformed scale labels (for example, 0, 1, and 2, showing the corresponding powers of 10). We have recently experienced the benefits and pitfalls of using log-scales as covid-19 dashboards displayed case count data on both the log and linear scale [Fagen-Ulmschneider, 2020, Financial Times, 2020]. In spring 2020, during the early stages of the coronavirus pandemic, there was a large magnitude discrepancy at a given time between different geographic regions (at all orders of magnitude - states and provinces as well as countries and continents). During this time, we saw the usefulness of log-scales showing case count curves for areas with few cases and areas with many cases within one chart. As the pandemic evolved, and the case counts were no longer spreading exponentially, graphs with linear scales seemed more effective at spotting early increases in case counts that signaled more localized outbreaks. This is only one recent example of a situation in which both log and linear scales are useful for showing different aspects of the same data; there are long histories of using log scales to display results in ecology, psychophysics, engineering, and physics [Munroe, 2013, Menge et al., 2018, Heckler et al., 2013].

Research suggests our perception and mapping of numbers to a number line is logarithmic at first, but transitions to a linear scale later in development, with formal mathematics education [Varshney and Sun, 2013, Siegler and Braithwaite, 2017, Dehaene et al., 2008]. This transition to linear scales occurs first in small numbers (e.g. 1-10) and then gradually expands to higher orders of magnitude; thus, the logarithmic intuition about numbers in children is often more noticeable on scales in the thousands to hundreds of thousands. If we percieve logarithmically by default, it is a natural (and presumably low-effort) way to display information and should be easy to read and understand/use. In fact, early studies explored the estimation and prediction of exponential growth, finding that growth is underestimated when presented both numerically and graphically but that numerical estimation is more accurate than graphical estimation for exponential curves. One way to improve estimation of increasing exponential trends is to provide immediate feedback to participants [Mackinnon and Wearing, 1991]. While prior contextual knowledge or experience with exponential growth does not improve estimation, instruction on exponential growth reduces the underestimation: participants adjust their initial starting value but not their perception of growth rate [Wagenaar and Sagaria, 1975, Jones, 1977].

Our inability to accurately predict exponential growth might also be addressed by log-transforming the data, however, this transformation introduces new complexities: most readers are not mathematically sophisticated enough to intuitively understand logarithmic math and translate that back into real-world effects. In Menge et al. [2018], ecologists were surveyed to determine how often ecologists encounter log-scaled data and how well ecologists understand log-scaled data when they see it in the literature. Participants were presented two relationships displayed on linear-linear scales, log-log scales with untransformed values, or log–log scales with log-transformed values. Menge et al. [2018] propose three types of misconceptions participants encountered when presented data on log-log scales: 'hand-hold fallacy', 'Zeno's zero fallacy', and 'watch out for curves fallacies'.

In order to provide a set of principles to guide design choices, we must evaluate these design choices through the use of graphical tests. These tests may take many forms: identifying differences in graphs, reading information off of a chart accurately, using data to make correct real-world decisions, and predicting the next few observations. All of these types of tests require different levels of use and manipulation of the information presented in the chart. To lay a foundation for future exploration of the use of log scales, we begin with the most fundamental ability to identify differences in charts: this does not require that participants understand exponential growth, identify log scales, or have any mathematical training. Instead, we are simply testing the change in perceptual sensitivity resulting from visualization choices. In Best et al. [2007], the authors explored whether discrimination between curve types is possible. They found that accuracy is higher when nonlinear trends presented (e.g. it's hard to say something is linear, but easy to say that it isn't) and that accuracy is higher with low additive variability.

A statistic is a numerical function which summarizes the data; by this definition, graphs are visual statistics. To evaluate a graph, we have to run our statistic throuugh a visual evaluation - a person. If two different methods of presenting data result in qualitatively different results when evaluated visually, then we can conclude that the visual statistics are significantly different. Recent graphical experiments have utilized statistical lineups to quantify the perception of graphical design choices[VanderPlas and Hofmann, 2017, Hofmann et al., 2012, Loy et al., 2016]. Statistical lineups provide an elegant way of combining perception and statistical hypothesis testing using graphical experiments [Wickham et al., 2010, Majumder et al., 2013, Vanderplas et al.]. 'Lineups' are named after the 'police lineup' of criminal investigations where witnesses are asked to identify the criminal from a set of individuals. Similarly, a statistical lineup is a plot consisting of smaller panels; the viewer is asked to identify the plot of the real data from among a set of decoy null plots. A statistical lineup typically consists of 20 panels - 1 target panel and 19 null panels (Figure 1). If the viewer can identify the target panel embeded within the set of null panels, this suggests that the real data is visually distinct from data generated under the null model. Crowd sourcing websites such as Amazon Mechanical Turk, Reddit, and Proflic allow us to collect responses from multiple viewers. In this paper, we use statistical lineups to test our ability to differentiate between exponentially increasing curves with differing growth rates, using linear and log scales.

2 Methodology

2.1 Data Generation

In this study, both the target and null data sets were generated by simulating data from an exponential model; the models differ in the parameters selected for the null and target panels. In order to guarantee the simulated data spans the same range of values, we implemented a range constraint of $y \in [10, 100]$ and a domain constraint of $x \in [0, 20]$ with $N = 50$ points randomly assigned throughout the domain and mapped to the y-axis using the exponential model with the selected parameters. These constraints provide some assurance that participants who select the target plot are doing so because of their visual perception differentiating between curvature or slope rather than different starting or ending values.

We simulated data based on a three-parameter exponential model with multiplicative errors:

$$y_i = \alpha \cdot e^{\beta \cdot x_i + \epsilon_i} + \theta \qquad (1)$$
$$\text{with } \epsilon_i \sim N(0, \sigma^2).$$

The parameters α and θ are adjusted based on β and σ^2 to guarantee the range and domain constraints are met. The model generated $N = 50$ points $(x_i, y_i), i = 1, ..., N$ where x and y have an increasing exponential relationship. The heuristic data generation procedure is provided in Appendix A.

The exponential model provides the base for this graphical experiment. We manipulate the midpoint, x_{mid}, and in turn the estimated parameters to control the amount of curvature present in the data and the error standard deviation, σ, to control the amount of deviation from the exponential curve. We selected three midpoints corresponding to difficulty levels easy (obvious curvature), medium (noticeable curvature), and hard (almost linear) along with a sensible choice of standard deviation, σ. The midpoints and standard deviation combinations were chosen using a method described in VanderPlas and Hofmann [2017]; additional details and final parameter values are available in Appendix B.

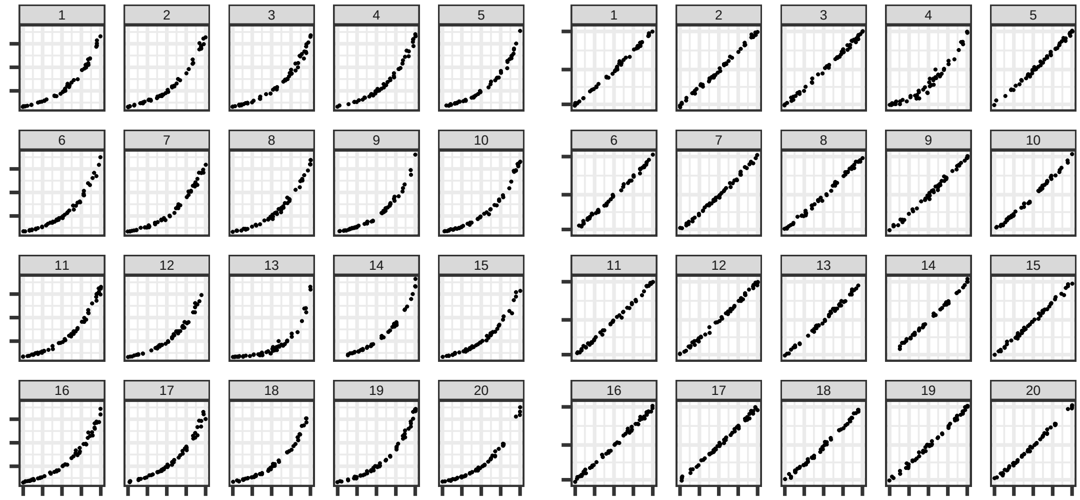

Figure 1: The plot on the left displays increasing exponential data on a linear scale with panel (2 x 5) + 3 as the target. The plot on the right displays increasing exponential data on the log scale with panel 2 x 2 as the target.

2.2 Lineup Setup

The lineup plots were generated by mapping simulating data corresponding to difficulty level A to a scatterplot while the null plots were generated by mapping simulated data corresponding to difficulty level B to a scatterplot. For example, a target plot with simulated data following an increasing exponential curve with obvious curvature is embeded within null plots with simulated data following an increasing exponential that is almost linear (i.e. Easy-Hard). By our constraints, the target plot and null plots will span a similar domain and range. There are a total of 6 (i.e. $3! \cdot 2!$) lineup parameter combinations. Two sets of each lineup parameter combination were simulated (total of 12 test data sets) and plotted on both the linear and the log scale (total of 24 test lineup plots). In addition, there are three parameter combinations which generate homogeneous "Rorschach" lineups, where all panels are from the same distribution. Each participant evaluated one of these lineups, but for simplicity, these evaluations are not described in this paper.

2.3 Study Design

Each participant was shown a total of thirteen lineup plots (twelve test lineup plots and one Rorschach lineup plot). Participants were randomly assigned one of the two replicate data sets for each of the six unique lineup parameter combinations. For each assigned test data set, the participant was shown the lineup plot corresponding to both the linear scale and the log scale. For the additional Rorschach lineup plot, participants were randomly assigned a one data set shown on either the linear or the log scale. The order of the thirteen lineup plots shown was randomized for each participant.

Participants above the age of majority were recruited from Reddit's Visualization and Sample Size communities. Since participants recruited on Reddit were not compensated for their time, most participants have an interest in data visualization research. Previous literature suggests that prior mathematical knowledge or experience with exponential data is not associated with the outcome of graphical experiments [VanderPlas and Hofmann, 2016]. Participants completed the experiment using a Shiny applet (`https://shiny.srvanderplas.com/log-study/`).

Participants were shown a series of lineup plots and asked to identify the plot that was most different from the others. On each plot, participants were asked to justify their choice and provide their level of confidence in their choice. The goal of this experimental task is to test an individuals ability to perceptually differentiate exponentially increasing data with differing rates of change on both the linear and log scale.

3 Results

Participant recruitment through Reddit occurred over the course of two weeks during which 58 individuals completed 518 unique test lineup evaluations. Participants who completed fewer than 6 lineup evaluations were removed from the study (17 participants, 41 evaluations). The final data set included a total of 41 participants and 477 lineup evaluations. Each plot was evaluated by between 18 and 28 individuals (Mean: 21.77, SD: 2.29). In

67% of the 477 lineup evaluations, participants correctly identified the target panel.

Target plot identification was analyzed using the Glimmix Procedure in SAS 9.4. Each lineup plot evaluated was assigned a value based on the participant response (correct = 1, not correct = 0). The binary response was analyzed using generalized linear mixed model following a binomial distribution with a logit link function following a row-column blocking design accounting for the variation due to participant and data set respectively. See model details and estimates in Appendix C

On both the log and linear scales, the highest accuracy occured in lineup plots where the target model and null model had large curvature differences (e.g. easy-hard and hard-easy). When comparing models that have slight curvature differences (e.g. Medium-Easy), there is a sacrifice in accuracy when displayed on the linear scale. It is worth noting that there is a more significant decrease in accuracy on the linear scale when comparing a target plot with a lower rate of change embedded in null plots with higher rates of change (e.g. Medium-Easy and Hard-Medium). This agrees with Best et al. [2007] whose results found that accuracy was higher when nonlinear trends were presented indicating that it is hard to say something is linear (i.e. something has less curvature), but easy to say that it isn't (i.e. something has more curvature). Overall, there is are no significant differences in accuracy between curvature combinations when data is presented on a log scale indicating participants were consistent in their success of identifying the target panel on the log scale. Figure 2 displays the estimated (log) odds ratio of successfully identifying the target panel on the log scale compared to the linear scale. We found the choice of scale has no impact if curvature differences are large. However, presenting data on the log scale makes us more sensitive to the the changes when there are only slight changes in curvature.

4 Discussion and Conclusion

The overall goal of this paper is to provide basic research to support the principles used to guide design decisions in scientific visualizations of exponential data. In this study, we explored the use of linear and log scales to determine whether our ability to notice differences in the data is different when different scales are used. Our results indicated that when there was a larger curvature difference, participants accurately differentiated between the two curves on both the linear and log scale. However, displaying increasing exponential data on a log scale improved the accuracy of differentiating between models with slight curvature differences. An exception occurs when identifying a plot with more curvature than the surrounding plots, supporting Best et al. [2007] whose results found that accuracy was higher when nonlinear trends were presented indicating that it is hard to say something is linear (i.e. something has less curvature), but easy to say that it isn't (i.e. something has more curvature).

Further experimentation is necessary to test an individual's ability to make predictions for exponentially increasing data. Previous literature suggests that we tend to underestimate predictions of exponentially increasing data.[Jones, 1979, 1977, Wagenaar and Timmers, 1978]. [Mosteller et al., 1981] designed and carried out an empirical investigation to explore properties of lines fitted by eye. The researchers found that students tended to fit the slope of the first principal component or major axis (the line that minimizes the sum of squares of perpendicular rather than vertical distances) and that students who overestimated the steepness of the slope were consistent in their overestimation throughout the entire study. Interestingly, the study found that individual-to-individual variability in slope and in intercept was near the standard error provided by least squares. A similar graphical task is used in the New York Times "You Draw It" page asking readers to test their knowledge by using their curser to estimate values of a certain topic under different political administrations or over different years [New York Times, 2017]. In addition to differentiation and prediction of exponentially increasing data, it is of interest to test an individuals' ability to translate a graph of exponentially increasing data into real value quantities and extend their estimations by making comparisons. [Friel et al., 2001] emphasize the importance of graph comprehension proposing that the graph construction plays a role in the ability to read and interpret graphs.

References

J. J. V. Bavel, K. Baicker, P. S. Boggio, V. Capraro, A. Cichocka, M. Cikara, M. J. Crockett, A. J. Crum, K. M. Douglas, J. N. Druckman, J. Drury, O. Dube, N. Ellemers, E. J. Finkel, J. H. Fowler, M. Gelfand, S. Han, S. A. Haslam, J. Jetten, S. Kitayama, D. Mobbs, L. E. Napper, D. J. Packer, G. Pennycook, E. Peters, R. E. Petty, D. G. Rand, S. D. Reicher, S. Schnall, A. Shariff, L. J. Skitka, S. S. Smith, C. R. Sunstein, N. Tabri, J. A. Tucker, S. v. d. Linden, P. v. Lange, K. A. Weeden, M. J. A. Wohl, J. Zaki, S. R. Zion, and R. Willer. Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, 4(5):460–471, May 2020. ISSN 2397-3374. doi: 10.1038/s41562-020-0884-z. URL `http://www.nature.com/articles/s41562-020-0884-z`.

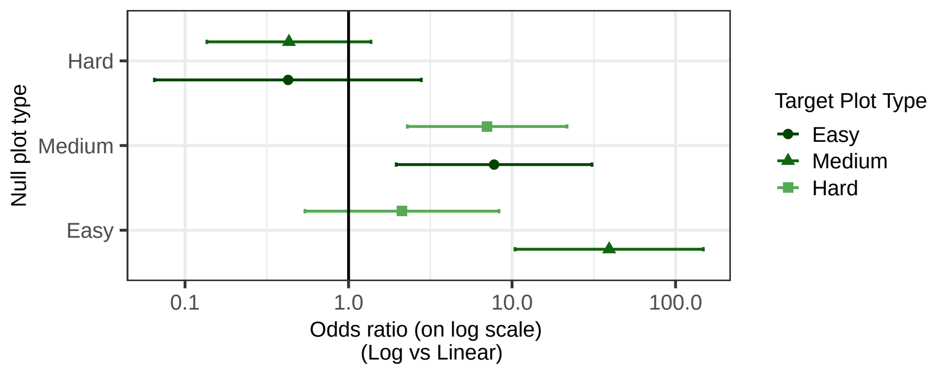

Figure 2: Odds Ratio's

L. A. Best, L. D. Smith, and D. A. Stubbs. Perception of Linear and Nonlinear Trends: Using Slope and Curvature Information to Make Trend Discriminations. *Perceptual and Motor Skills*, 104(3):707–721, June 2007. ISSN 0031-5125. doi: 10.2466/pms.104.3.707-721. URL `https://doi.org/10.2466/pms.104.3.707-721`. Publisher: SAGE Publications Inc.

L. Charlotte. You've informed the public with visualizations about the coronavirus. thank you., Jul 2020. URL `https://blog.datawrapper.de/datawrapper-effect-corona/`.

S. Dehaene, V. Izard, E. Spelke, and P. Pica. Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures. *Science*, 320(5880):1217–1220, May 2008. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1156540. URL `https://science.sciencemag.org/content/320/5880/1217`. 00651 Publisher: American Association for the Advancement of Science Section: Report.

W. Fagen-Ulmschneider. 91-divoc, 2020. URL `http://91-divoc.com/pages/covid-visualization/`.

Financial Times. Coronavirus tracked: has the epidemic peaked near you?, Dec 2020. URL `https://ig.ft.com/coronavirus-chart/?areas=eur`.

S. N. Friel, F. R. Curcio, and G. W. Bright. Making Sense of Graphs: Critical Factors Influencing Comprehension and Instructional Implications. *Journal for Research in Mathematics Education*, 32(2):124, Mar. 2001. ISSN 00218251. doi: 10.2307/749671. URL `https://www.jstor.org/stable/749671?origin=crossref`.

A. F. Heckler, B. Mikula, and R. Rosenblatt. Student accuracy in reading logarithmic plots: The problem and how to fix it. In *2013 IEEE Frontiers in Education Conference (FIE)*, pages 1066–1071, Oct. 2013. doi: 10.1109/FIE.2013.6684990. ISSN: 2377-634X.

H. Hofmann, L. Follett, M. Majumder, and D. Cook. Graphical Tests for Power Comparison of Competing Designs. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2441–2448, Dec. 2012. ISSN 1077-2626. doi: 10.1109/TVCG.2012.230. URL `http://ieeexplore.ieee.org/document/6327249/`.

G. V. Jones. Polynomial perception of exponential growth. *Perception & Psychophysics*, 21(2):197–198, Mar. 1977. ISSN 0031-5117, 1532-5962. doi: 10.3758/BF03198726. URL `http://link.springer.com/10.3758/BF03198726`.

G. V. Jones. A generalized polynomial model for perception of exponential series. *Perception & Psychophysics*, 25(3):232–234, May 1979. ISSN 0031-5117, 1532-5962. doi: 10.3758/BF03202992. URL `http://link.springer.com/10.3758/BF03202992`.

A. Loy, L. Follett, and H. Hofmann. Variations of Q-Q Plots: The Power of Our Eyes! *The American Statistician*, 70(2):202–214, Apr. 2016. ISSN 0003-1305, 1537-2731. doi: 10.1080/00031305.2015.1077728. URL `https://www.tandfonline.com/doi/full/10.1080/00031305.2015.1077728`. 00000.

A. J. Mackinnon and A. J. Wearing. Feedback and the forecasting of exponential change. *Acta Psychologica*, 76(2):177–191, Apr. 1991. ISSN 00016918. doi: 10.1016/0001-6918(91)90045-2. URL `https://linkinghub.elsevier.com/retrieve/pii/0001691891900452`.

M. Majumder, H. Hofmann, and D. Cook. Validation of Visual Statistical Inference, Applied to Linear Models. *Journal of the American Statistical Association*, 108(503):942–956, Sept. 2013. ISSN 0162-1459, 1537-274X. doi: 10.1080/01621459.2013.

808157. URL http://www.tandfonline.com/doi/abs/10.1080/01621459.2013.808157.

D. N. L. Menge, A. C. MacPherson, T. A. Bytnerowicz, A. W. Quebbeman, N. B. Schwartz, B. N. Taylor, and A. A. Wolf. Logarithmic scales in ecological data presentation may cause misinterpretation. *Nature Ecology & Evolution*, 2(9):1393–1402, Sept. 2018. ISSN 2397-334X. doi: 10.1038/s41559-018-0610-7. URL http://www.nature.com/articles/s41559-018-0610-7.

F. Mosteller, A. F. Siegel, E. Trapido, and C. Youtz. Eye Fitting Straight Lines. *The American Statistician*, 35(3):150–152, Aug. 1981. ISSN 0003-1305, 1537-2731. doi: 10.1080/00031305.1981.10479335. URL http://www.tandfonline.com/doi/abs/10.1080/00031305.1981.10479335.

R. Munroe. Log scale, Jan 2013. URL https://xkcd.com/1162/.

New York Times. Ny times you draw it charts, Jan 2017. URL https://presentyourstory.com/ny-times-you-draw-it-charts/.

A. Romano, C. Sotis, G. Dominioni, and S. Guidi. The Scale of COVID-19 Graphs Affects Understanding, Attitudes, and Policy Preferences. SSRN Scholarly Paper ID 3588511, Social Science Research Network, Rochester, NY, Apr. 2020. URL https://papers.ssrn.com/abstract=3588511.

L. C. Rost. You've informed the public with visualizations about the coronavirus. thank you., Jul 2020. URL https://blog.datawrapper.de/datawrapper-effect-corona/.

R. S. Siegler and D. W. Braithwaite. Numerical Development. *Annual Review of Psychology*, 68(1):187–213, Jan. 2017. ISSN 0066-4308, 1545-2085. doi: 10.1146/annurev-psych-010416-044101. URL http://www.annualreviews.org/doi/10.1146/annurev-psych-010416-044101.

A. Unwin. Why is Data Visualization Important? What is Important in Data Visualization? *Harvard Data Science Review*, Jan. 2020. doi: 10.1162/99608f92.8ae4d525. URL https://hdsr.mitpress.mit.edu/pub/zok97i7p.

S. VanderPlas and H. Hofmann. Spatial Reasoning and Data Displays. *IEEE Transactions on Visualization & Computer Graphics*, 22(1):459–468, Jan. 2016. ISSN 1077-2626. doi: 10.1109/TVCG.2015.2469125. URL doi.ieeecomputersociety.org/10.1109/TVCG.2015.2469125.

S. VanderPlas and H. Hofmann. Clusters Beat Trend!? Testing Feature Hierarchy in Statistical Graphics. *Journal of Computational and Graphical Statistics*, 26(2):231–242, Apr. 2017. ISSN 1061-8600, 1537-2715. doi: 10.1080/10618600.2016.1209116. URL https://www.tandfonline.com/doi/full/10.1080/10618600.2016.1209116.

S. Vanderplas, C. Röttger, D. Cook, and H. Hofmann. Statistical significance calculations for scenarios in visual inference. *Stat*, page e337. URL https://doi.org/10.1002/sta4.337.

S. Vanderplas, D. Cook, and H. Hofmann. Testing statistical charts: What makes a good graph? *Annual Review of Statistics and Its Application*, 7:61–88, 2020.

L. R. Varshney and J. Z. Sun. Why do we perceive logarithmically? *Significance*, 10(1):28–31, Feb. 2013. ISSN 17409705. doi: 10.1111/j.1740-9713.2013.00636.x. URL http://doi.wiley.com/10.1111/j.1740-9713.2013.00636.x.

W. A. Wagenaar and S. D. Sagaria. Misperception of exponential growth. *Perception & Psychophysics*, 18(6):416–422, Nov. 1975. ISSN 0031-5117, 1532-5962. doi: 10.3758/BF03204114. URL http://link.springer.com/10.3758/BF03204114.

W. A. Wagenaar and H. Timmers. Extrapolation of exponential time series is not enhanced by having more data points. *Perception & Psychophysics*, 24(2):182–184, Mar. 1978. ISSN 0031-5117, 1532-5962. doi: 10.3758/BF03199548. URL http://link.springer.com/10.3758/BF03199548.

H. Wickham, D. Cook, H. Hofmann, and A. Buja. Graphical inference for infovis. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):973–979, 2010.

A Data Generation Procedure

Algorithm 2.1.1: Paremeter Estimation

Input Parameters: domain $x \in [0, 20]$, range $y \in [10, 100]$, midpoint x_{mid}.

Output: estimated model parameters $\hat{\alpha}, \hat{\beta}, \hat{\theta}$

1. Determine the $y = -x$ line scaled to fit the assigned domain and range.

2. Map the values $x_{mid} - 0.1$ and $x_{mid} + 0.1$ to the $y = -x$ line for two additional points.

3. From the set points (x_k, y_k) for $k = 1, 2, 3, 4$, obtain the coefficients from the linear model $\ln(y_k) = b_0 + b_1 x_k$ to obtain starting values - $\alpha_0 = e^{b_0}, \beta_0 = b_1, \theta_0 = 0.5 \cdot \min(y)$

4. Using the `nls()` function from the `stats` package in Rstudio and the starting parameter values - $\alpha_0, \beta_0, \theta_0$ - fit the nonlinear model, $y_k = \alpha \cdot e^{\beta \cdot x_k} + \theta$ to obtain estimated parameter values - $\hat{\alpha}, \hat{\beta}, \hat{\theta}$.

Algorithm 2.1.2: Exponential Simulation

Input Paremeters: sample size $N = 50$, estimated parameters $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\theta}$, σ standard deviation from the exponential curve.

Output Parameters: N points, in the form of vectors $\mathbf{x}$ and $\mathbf{y}$.

1. Generate $\tilde{x}_j, j = 1, ..., N \cdot \frac{3}{4}$ as a sequence of evenly spaced points in $[0, 20]$. This ensures the full domain of x is used, fulfilling the constraints of spanning the same domain and range for each parameter combination.

2. Obtain $\tilde{x}_i, i = 1, ...N$ by sampling $N = 50$ values from the set of $\tilde{x}_j$ values. This gaurantees some variability and potential clustring in the exponential growth curve disrupting the perception due to continuity of points.

3. Obtain the final x_i values by jittering $\tilde{x}_i$.

4. Calculate $\tilde{\alpha} = \frac{\hat{\alpha}}{e^{\sigma^2/2}}$. This ensures that the range of simulated values for different standard devaition parameters has an equal expected value for a given rate of change due to the non-constant variance across the domain.

5. Generate $y_i = \tilde{\alpha} \cdot e^{\hat{\beta} x_i + e_i} + \hat{\theta}$ where $e_i \sim N(0, \sigma^2)$.

B Parameter Selection

For each level of difficulty, we simulated 1000 data sets of (x_{ij}, y_{ij}) points for $i = 1, ..., 50$ and $j = 1...10$. Each generated x_i point from *Algorithm 2.1.2* was replicated 10 times. Then the lack of fit statistic (LOF) was computed for each simulated data set by calculating the deviation of the data from a linear line. Plotting the density curves of the LOF statistics for each level of difficulty choice allows us to evaluate the ability of differentiating between the difficulty levels and thus detecting the target plot. In Figure 3, we can see the densities of each of the three difficulty levels. While the LOF statistic provides us a numerical value for discriminating between the difficulty levels, we cannot directly relate this to the perceptual discriminability; it serves primarily as an approximation to ensure that we are testing parameters at several distinct levels of difficulty.

Final parameter estimates are shown in Table 1.

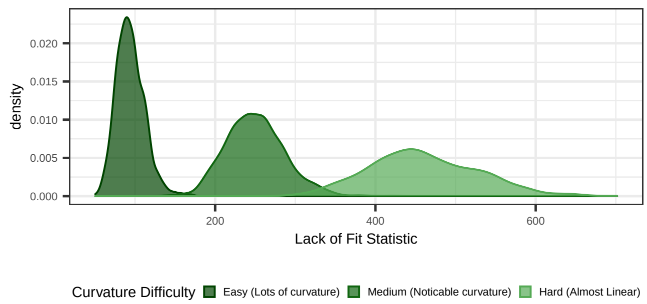

Figure 3: Density plot of the lack of fit statistic showing separation of difficulty levels: obvious curvature, noticable curvature, and almost linear.

Table 1: Final parameters used to simulate exponential data for each of the three diffculty levels: Easy (Obvious curvature), Medium (noticable curvature) and Hard (almost linear)

	x_{mid}	$\hat{\alpha}$	$\tilde{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\sigma}$
Easy	14.5	0.91	0.88	0.23	9.10	0.25
Medium	13.0	6.86	6.82	0.13	3.14	0.12
Hard	11.5	37.26	37.22	0.06	-27.26	0.05

C Model Details

Target plot identification was analyzed using the Glimmix Procedure in SAS 9.4. Each lineup plot evaluated was assigned a value based on the participant response (correct = 1, not correct = 0). Define Y_{ijkl} to be the event that participant l correctly identifies the target plot for data set k with curvature j plotted on scale i. The binary response was analyzed using generalized linear mixed model following a binomial distribution with a logit link function following a row-column blocking design accounting for the

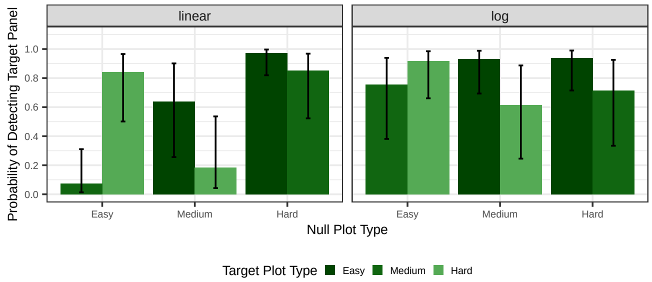

Figure 4: Least Squares Means

Table 2: Type III Tests for Fixed Effects

Fixed Effect	F	DF	P-value
Curvature	3.66	5 , 419	0.003
Scale	14.89	1 , 419	0.0001
Curvature x Scale	6.58	5 , 419	$<$0.0001

variation due to participant and data set respectively as shown in the Equation (2).

$$\text{logit } P(Y_{ijk}) = \eta + \delta_i + \gamma_j + \delta\gamma_{ij} + s_l + d_k \qquad (2)$$

where

- η is the beaseline average probability of selecting the target plot.
- δ_i is the effect of the log/linear scale.
- γ_j is the effect of the curvature combination.
- $\delta\gamma_{ij}$is the two-way interaction effect of the scale and curvature.
- $s_l \sim N(0, \sigma^2_{\text{participant}})$, random effect for participant characteristics.
- $d_k \sim N(0, \sigma^2_{\text{data}})$, random effect for data specific characteristics.

We assume that random effects for data set and participant are independent. See least squares means estimated probabilities in Figure 4.

Type III tests for fixed effects shown in Table 2 indicate a significant interaction between the curvature combination and scale. Variance due to participant and data set were estimated to be $\sigma^2_{participant} = 2.13$ (s.e. 0.74) and $\sigma^2_{dataset} = 0.92$ (s.e. 0.70).