HUMAN PERCEPTION OF EXPONENTIALLY INCREASING DATA DISPLAYED ON A LOG SCALE

by

Emily Anna Robinson

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

For the Degree of Doctor of Philosophy

Major: Statistics

Under the Supervision of Professors Susan VanderPlas and Reka Howard

Lincoln, Nebraska

August, 2022

HUMAN PERCEPTION OF EXPONENTIALLY INCREASING DATA DISPLAYED ON A LOG SCALE

Emily Anna Robinson, Ph.D.

University of Nebraska, 2022

Advisers: Susan VanderPlas and Reka Howard

Log scales are often used to display data over several orders of magnitude within one graph. During the COVID-19 pandemic, we have seen both the benefits and the pitfalls of using log scales to display case counts. Three graphical experimental tasks were conducted to evaluate the impact our choice of scale has on human perception of exponentially increasing trends. The first experiment evaluates whether our ability to perceptually notice differences in exponentially increasing trends is impacted by the choice of scale. We conducted a visual inference experiment in which participants were shown a series of lineup plots (consisting of 19 null panels and 1 target panel) and asked to identify the panel that was most different from the others. Our results indicated that when there was a large difference in curvature between the target plot and null plots, the choice of scale had no impact and participants accurately differentiated between the two curves on both the linear and log scale. However, displaying exponentially increasing data on a log scale improved the accuracy of differentiating between models with slight curvature differences. An exception occurred when identifying a plot with curvature embedded in surrounding plots closely relating to a linear trend, indicating that it is easy to identify a curve in a group of lines but much harder to identify a line in a group of curves. The use of visual inference to

identify these guidelines suggests that there are *perceptual* advantages to log scales when differences are subtle. Our other experimental tasks focus on determining whether there are cognitive disadvantages to log scales: do log scales make it harder to make use of graphical information? We conducted a graphical task similar to the New York Times "You Draw It" page to test an individual's ability to use and make predictions for exponentially increasing data. We asked participants to draw a line using their computer mouse through the increasing exponential trend shown on both scales. In addition to differentiation and prediction of exponentially increasing data, we conduct an experimental task to test an individuals' ability to translate a graph of exponentially increasing data into real value quantities and extend their estimations by making comparisons. The results of our experimental tasks allow us to provide guidelines for readers to actively choose which of many possible graphics to draw, according to some set of design choices, to ensure that our charts are effective. *(399 words; 350 word limit)*

Thank you to all my people!

Table of Contents

List of Figures

List of Tables

Chapter 1

Literature Review

Editing text colors: Emily's editing color. Emily may also use mostly black text as well. Susan's editing color. Reka's editing color.

- Introduction

With the advancement of technology, access to data is more readily available more than ever before. With this abundance of data comes the responsibility to make sense out of the numbers in order to convey the message clearly. Graphics are a useful tool for displaying and communicating information (Vanderplas, Cook, & Hofmann, 2020). Researchers include graphics to communicate their results in scientific publications and news sources rely on graphics to convey news stories to the public [CITE ME].

- Motivation (COVID-19)

    - Charlotte (2020)

    - Romano, Sotis, Dominioni, & Guidi (2020)

    - Rost (2020)

    - Bavel et al. (2020)

– Fagen-Ulmschneider (2020)

– Financial Times (2020)

At the beginning of the SARS-NCOV-2 pandemic (COVID-19), 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 share information with the public about the progression of the pandemic (Romano, Sotis, Dominioni, & Guidi, 2020). People began seeking out graphical displays of COVID-19 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, as well as facilitated communication with the public to increase compliance (Bavel et al., 2020).

- General Sensation and Perception
- General Graphics Research

  – Vanderplas, Cook, & Hofmann (2020)

  – Unwin (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).

- Logrithms

- Menge et al. (2018)

- Munroe (2013)

- Heckler, Mikula, & Rosenblatt (2013)

- Varshney & Sun (2013)

- Siegler & Braithwaite (2017)

- Dehaene, Izard, Spelke, & Pica (2008)

- Visual Inference through Lineups

  - Best, Smith, & Stubbs (2007)

  - Susan VanderPlas & Hofmann (2017)

  - Hofmann, Follett, Majumder, & Cook (2012)

  - Loy, Follett, & Hofmann (2016)

  - Wickham, Cook, Hofmann, & Buja (2010)

  - Majumder, Hofmann, & Cook (2013)

  - Vanderplas, Röttger, Cook, & Hofmann (2020)

- Prediction with You Draw It

  - William A. Wagenaar & Sagaria (1975)

  - Mackinnon & Wearing (1991)

  - Jones (1977)

  - Jones (1979)

  - W. A. Wagenaar & Timmers (1978)

  - Timmers & Wagenaar (1977)

  - Mosteller, Siegel, Trapido, & Youtz (1981)

  - Finney (1951)

- Numerical Translation and Estimation

Text below copied from JSM 2021 Student Paper

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 Financial Times (2020). In spring 2020, during the early stages of the COVID-19 pandemic, there were large magnitude discrepancies in case counts at a given time point between different geographic regions (e.g. states and provinces as well as countries and continents). During this time, we saw the usefulness of log scale transformations 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 Heckler, Mikula, & Rosenblatt (2013). In fact, research shows that we do not perceive exponential growth accurately (whether information is presented in graphical or tabular form): we systematically under-estimate exponential growth.

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 Dehaene, Izard, Spelke, & Pica (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 perceive 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 (William A. Wagenaar & Sagaria, 1975). One way to improve estimation of increasing exponential trends is to provide immediate feedback to participants about the accuracy of their current predictions (Mackinnon & 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 Jones (1977). Estimation was shown to improve when subjects were

presented with decreasing exponential functions (Timmers & Wagenaar, 1977). Jones (1977),Jones (1979) and W. A. Wagenaar & Timmers (1978) propose competing polynomial models for the perception and extrapolation of exponential series. It seems that estimation is a two-stage process: first, we identify the type of curve and direction and then, we use that information for prediction (Best, Smith, & Stubbs, 2007).

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.' These misconceptions are a result of linear extrapolation assuming that a line in log-log space represents a line instead of the power law in linear-linear space. The study found that in each of these scenarios, participants were confident in their incorrect responses, indicating incorrect knowledge rather than a lack of knowledge.

The 'hand-hold falacy' stems from the misconception that steeper slopes in log-log relationships are steeper slopes in linear-linear space. In fact, it is not

only the slope that matters, but also the intercept and the location on the horizontal axis since a line in log-log space represents a power law in linear-linear space (i.e. linear extraploation). Emerging from 'Zeno's zero fallacy' is the misconception that positively sloped lines in log-log space can imply a non-zero value of y when x is zero. This is never true as postively sloped lines in log-log space actually imply that y = 0 when x = 0. This misconception again is a result of linear extrapolation assuming that a line in log-log space represents a line instead of the power law in linear-linear space. The last misconception, 'watch out for curves fallacies' encompases three faults: (1) lines in log-log space are lines in linear-linear space, (2) lines in log-log space curve upward in linear-linear space, and (3) curves in log-log space have the same curvature in linear-linear space. Linear extrapolation is again responsible for the first and third faults while the second fault is a result of error in thinking that log-log lines represent power laws (which are exponential relationships), and all exponential relationships curve upward; this is only true when the log-log slope is greater than one. Menge et al. (2018) found that in each of these scenarios, participants were confident in their incorrect responses, indicating incorrect knowledge rather than a lack of knowledge.

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, or 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.

Chapter 2

Perception through lineups

2.1 Introduction

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,
Smith, & Stubbs (2007), the authors explored whether discrimination
between curve types is possible. They found that accuracy of identifying the
correct curve type is higher when nonlinear trends were presented (e.g. it's
hard to say something is linear, but easy to say that it is not linear) 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 through 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 choicesLoy, Follett, & Hofmann (2016). Statistical lineups provide an elegant way of combining perception and statistical hypothesis testing using graphical experiments Vanderplas, Röttger, Cook, & Hofmann (2020). '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 2.1). If the viewer can identify the target panel embedded 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 Prolific 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 levels of curvature, using linear and log scales. Vanderplas, Röttger, Cook, & Hofmann (2020) provides an approach for calculating visual p-values utilizing a 'rorschach' lineup which consists solely of null panels.

2.2   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 domain and range of values, we implemented
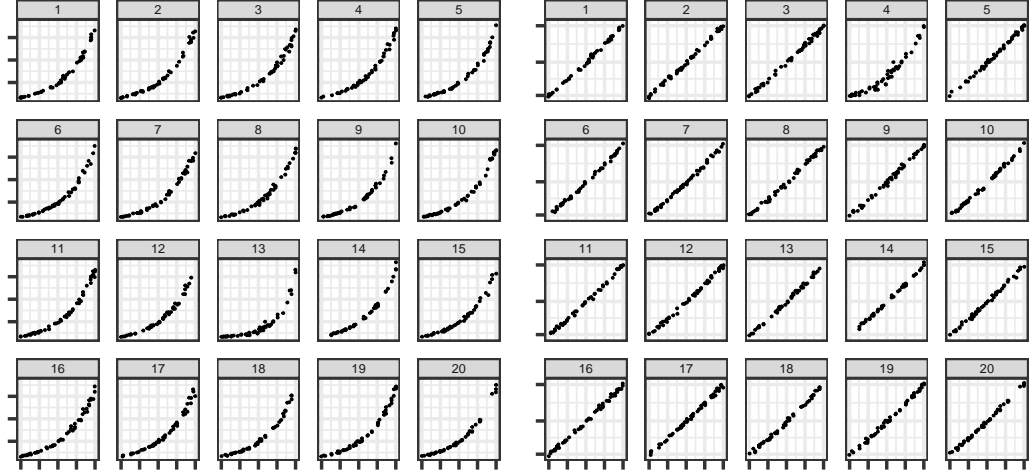
Figure 2.1: Lineup example

a domain constraint of $x \in [0, 20]$ and a range constraint of $y \in [10, 100]$ 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 growth rate rather than different starting or ending values.

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

$$y_i = \alpha \cdot e^{\beta \cdot x_i + \epsilon_i} + \theta \tag{2.1}$$

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 described below:

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 $e_i \sim N(0, \sigma^2)$.

2.3 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 2.2, 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.
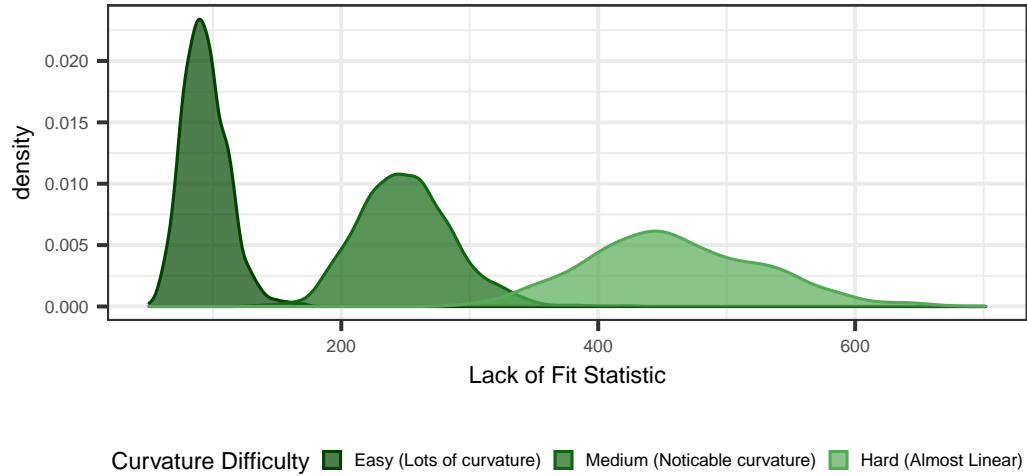
Figure 2.2: Lack of fit statistic density curves

Table 2.1: Lineup data simulation final parameters

	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

Final parameter estimates are shown in Table 2.1.

2.4   Lineup Setup

Lineup plots were generated by mapping one simulated data set corresponding to difficulty level A to a scatter plot to be identified as the target plot while multiple simulated data sets corresponding to difficulty level B were individually mapped to scatter plots for the null plots. For example, a target plot with simulated data following an increasing exponential curve with obvious curvature is embedded within null plots with simulated data following an increasing exponential trend that is almost linear (i.e. Hard Null - Easy Target). By our constraints, the target plot and null

plots will span a similar domain and range. There are a total of six (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 scale 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.

2.5 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 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 (S. VanderPlas & 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 trends with differing levels of curvature on both the linear and log scale.

2.6 Results

Participant recruitment through Reddit occurred over the course of two weeks during which 58 individuals completed 518 unique test lineup evaluations. Previous studies have found that results do not differ on lineup-related tasks between Reddit and e.g. Amazon Mechanical Turk (Susan VanderPlas & Hofmann, 2017). 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 a generalized linear mixed model following a binomial distribution with a logit link function following a row-column blocking design to account for the variation due to participant and data set respectively. See model details

and estimates in Appendix Reference.

On both the log and linear scales, the highest accuracy occurred in lineup plots where the target model and null model had large curvature differences (Easy Null - Hard Target; Hard Null - Easy Target). There is a decrease in accuracy on the linear scale when comparing a target plot with less curvature to null plots with more curvature (Easy Null - Medium Target; Medium Null - Hard Target). Best, Smith, & Stubbs (2007) found that accuracy of identifying the correct curve type 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 is not linear; our results concur with this observation. Overall, there 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 2.3 displays the estimated (log) odds ratio of successfully identifying the target panel on the log scale compared to the linear scale. The choice of scale has no impact if curvature differences are large (Hard Null - Easy Target; Easy Null - Hard Target). However, presenting data on the log scale makes us more sensitive to slight changes in curvature (Medium Null - Easy Target; Medium Null - Hard Target; Easy Null - Medium Target). An exception occurs when identifying a plot with curvature embedded in null plots close to a linear trend (Hard Null - Medium Target), again supporting the claim that it is easy to identify a curve in a bunch of lines but much harder to identify a line in a bunch of curves (Best, Smith, & Stubbs, 2007).
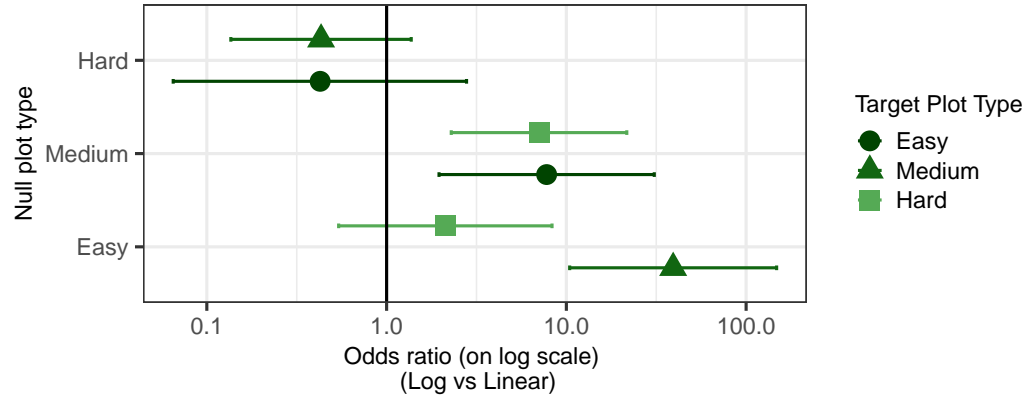
Figure 2.3: Lineups log(odds) results

2.7 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 explore the use of linear and log scales to determine whether our ability to notice differences in exponentially increasing trends is impacted by the choice of scale. Our results indicated that when there was a large difference in curvature between the target plot and null plots, the choice of scale had no impact and participants accurately differentiated between the two curves on both the linear and log scale. However, displaying exponentially increasing data on a log scale improved the accuracy of differentiating between models with slight curvature differences. An exception occurred when identifying a plot with curvature embedded in surrounding plots closely relating to a linear trend, indicating that it is easy to identify a curve in a group of lines but much harder to identify a line in a group of curves. The use of visual inference to identify these guidelines suggests that there are *perceptual* advantages to log scales when differences are subtle. What remains to be seen is whether there are

cognitive disadvantages to log scales: do log scales make it harder to make use of graphical information?

Chapter 3

Prediction with you draw it

Chapter 4

Numerical Translation and Estimation

Chapter 5

Conclusion

Appendix A

The First Appendix

Appendix B

The Second Appendix

References

10 Bavel, J. J. V., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., … Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, *4*(5), 460–471. http://doi.org/10.1038/s41562-020-0884-z

Best, L. A., Smith, L. D., & Stubbs, D. A. (2007). Perception of Linear and Nonlinear Trends: Using Slope and Curvature Information to Make Trend Discriminations. *Perceptual and Motor Skills*, *104*(3), 707–721. http://doi.org/10.2466/pms.104.3.707-721

Charlotte, L. (2020, July). You've informed the public with visualizations about the coronavirus. Thank you. *Chartable.* Retrieved from https://blog.datawrapper.de/datawrapper-effect-corona/

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures. *Science*, *320*(5880), 1217–1220. http://doi.org/10.1126/science.1156540

Fagen-Ulmschneider, W. (2020). 91-DIVOC. *An interactive visualization of the exponential spread of COVID-19.* Retrieved from http://91-divoc.com/pages/covid-visualization/

Financial Times. (2020, December). Coronavirus tracked: Has the epidemic peaked near you? *Coronavirus chart: see how your country compares.* Financial Times. Retrieved from https://ig.ft.com/coronavirus-chart/?areas=eur

Finney, D. (1951). Subjective judgment in statistical analysis: An experimental study. *Journal of the Royal Statistical Society: Series B (Methodological)*, *13*(2), 284–297.

Heckler, A. F., Mikula, B., & Rosenblatt, R. (2013). Student accuracy in reading logarithmic plots: The problem and how to fix it. In *2013 IEEE Frontiers in Education Conference (FIE)* (pp. 1066–1071). http://doi.org/10.1109/FIE.2013.6684990

Hofmann, H., Follett, L., Majumder, M., & Cook, D. (2012). Graphical Tests for Power Comparison of Competing Designs. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2441–2448. http://doi.org/10.1109/TVCG.2012.230

Jones, G. V. (1977). Polynomial perception of exponential growth. *Perception & Psychophysics*, *21*(2), 197–198. http://doi.org/10.3758/BF03198726

Jones, G. V. (1979). A generalized polynomial model for perception of exponential series. *Perception & Psychophysics*, *25*(3), 232–234. http://doi.org/10.3758/BF03202992

Loy, A., Follett, L., & Hofmann, H. (2016). Variations of Q-Q Plots: The Power of Our Eyes! *The American Statistician*, *70*(2), 202–214. http://doi.org/10.1080/00031305.2015.1077728

Mackinnon, A. J., & Wearing, A. J. (1991). Feedback and the forecasting of exponential change. *Acta Psychologica*, *76*(2), 177–191. http://doi.org/10.1016/0001-6918(91)90045-2

Majumder, M., Hofmann, H., & Cook, D. (2013). Validation of Visual Statistical Inference, Applied to Linear Models. *Journal of the American Statistical Association*, *108*(503), 942–956. http://doi.org/10.1080/01621459.2013.808157

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

Mosteller, F., Siegel, A. F., Trapido, E., & Youtz, C. (1981). Eye Fitting Straight Lines. *The American Statistician*, *35*(3), 150–152. http://doi.org/10.1080/00031305.1981.10479335

Munroe, R. (2013, January). Log scale. *xkcd*. Retrieved from https://xkcd.com/1162/

Romano, A., Sotis, C., Dominioni, G., & Guidi, S. (2020). *The Scale of COVID-19 Graphs Affects Understanding, Attitudes, and Policy Preferences* ({SSRN} {Scholarly} {Paper} No. ID 3588511). Rochester, NY: Social Science Research Network.

Rost, L. C. (2020, July). You've informed the public with visualizations about the coronavirus. Thank you. *Chartable*. Datawrapper. Retrieved from

https://blog.datawrapper.de/datawrapper-effect-corona/

Siegler, R. S., & Braithwaite, D. W. (2017). Numerical Development. *Annual Review of Psychology*, *68*(1), 187–213. http://doi.org/10.1146/annurev-psych-010416-044101

Timmers, H., & Wagenaar, W. A. (1977). Inverse statistics and misperception of exponential growth. *Perception & Psychophysics*, *21*(6), 558–562. http://doi.org/10.3758/BF03198737

Unwin, A. (2020). Why is Data Visualization Important? What is Important in Data Visualization? *Harvard Data Science Review.* http://doi.org/10.1162/99608f92.8ae4d525

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

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

VanderPlas, Susan, & Hofmann, H. (2017). Clusters Beat Trend!? Testing Feature Hierarchy in Statistical Graphics. *Journal of Computational and Graphical Statistics*, *26*(2), 231–242. http://doi.org/10.1080/10618600.2016.1209116

Vanderplas, S., Röttger, C., Cook, D., & Hofmann, H. (2020). Statistical significance calculations for scenarios in visual inference. *Stat.* http://doi.org/doi.org/10.1002/sta4.337

Varshney, L. R., & Sun, J. Z. (2013). Why do we perceive logarithmically? *Significance*, *10*(1), 28–31. http://doi.org/10.1111/j.1740-9713.2013.00636.x

Wagenaar, William A., & Sagaria, S. D. (1975). Misperception of exponential growth. *Perception & Psychophysics*, *18*(6), 416–422. http://doi.org/10.3758/BF03204114

Wagenaar, W. A., & Timmers, H. (1978). Extrapolation of exponential time series is not enhanced by having more data points. *Perception & Psychophysics*, *24*(2), 182–184. http://doi.org/10.3758/BF03199548

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