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

Emily A. Robinson 1
Department of Statistics, California Polytechnic State University - San Luis Obispo
and
Reka Howard 2
Department of Statistics, University of Nebraska - Lincoln
and
Susan VanderPlas 3
Department of Statistics, University of Nebraska - Lincoln

May 16, 2023

Abstract

(200 or fewer words) Logarithmic transformations are a standard solution to displaying data that spans several magnitudes within a single graph. This paper investigates the impact of log scales on perceptual sensitivity through a visual inference experiment using statistical lineups. Our study evaluated participant's ability to detect differences between exponentially increasing data, characterized by varying levels of curvature, using both linear and logarithmic scales. Participants were presented with a series of plots and asked to identify the panel that appeared most different from the others. Due to the choice of scale altering the contextual appearance of the data, the results revealed slight perceptual advantages for both scales depending on the curvatures of the compared data. This study serves as the initial part of a three-paper series dedicated to understanding the perceptual and cognitive implications of using logarithmic scales for visualizing exponentially increasing data. Subsequent papers will investigate whether using log scales presents cognitive disadvantages, such as making it harder to utilize graphical information. Furthermore, these studies serve as an example of multi-modal graphical testing, examining different levels of engagement and interaction with graphics to establish nuanced and specific guidelines for graphical design.

Keywords: log scales, visual inference, graphical testing

1 Introduction

Effective communication of data is critical in influencing people's opinions and actions. This consideration was particularly true during the COVID-19 pandemic, where data visualizations and dashboards were vital in informing the public and policymakers about the outbreak's status. Local governments relied on graphics to inform their decisions about shutdowns and mask mandates, while residents were presented with data visualizations to encourage compliance with these regulations. A major issue designers encountered when creating COVID-19 plots was how to display data from a wide range of values (Fagen-Ulmschneider 2020, Burn-Murdoch et al. 2020). When faced with data that spans several orders of magnitude, we must decide whether to show the data on its original scale (compressing the smaller magnitudes into a relatively small area) or to transform the scale and alter the contextual appearance of the data. Log axis transformations have emerged as a standard solution to this challenge, as they allow for the display of data over several orders of magnitude within a single graph.

Exponential data are one such example of a function that compresses smaller magnitudes into a smaller area; Fig. 1 presents hard drive capacity over the past forty years on both the linear and log scales to demonstrate the usefulness of log scales when dealing with data spanning multiple magnitudes. Logarithms facilitate the conversion of multiplicative relationships (displaying 1 & 10 with a distance of 10 units apart and displaying 10 & 100 with a distance of 90 units apart) to additive relationships (displaying 1 & 10 and 10 & 100 an equal distance apart), highlighting proportional relationships and linearizing power functions (Menge et al. 2018). Logarithms also have practical applications, simplifying the computation of small numbers such as likelihoods and transforming data to conform to statistical assumptions. Although log scales have a long history of use in fields such as ecology, psychophysics, engineering, and physics (Heckler et al. 2013, Waddell 2005), there is still a need to understand the implications of their use and provide best practices for their implementation.

In this paper, we evaluate the benefits and drawbacks of using log scales and examine their impact on perceptual sensitivity by conducting a visual inference experiment using statistical lineups (Buja et al. 2009). The experiment focused on a participant's ability to

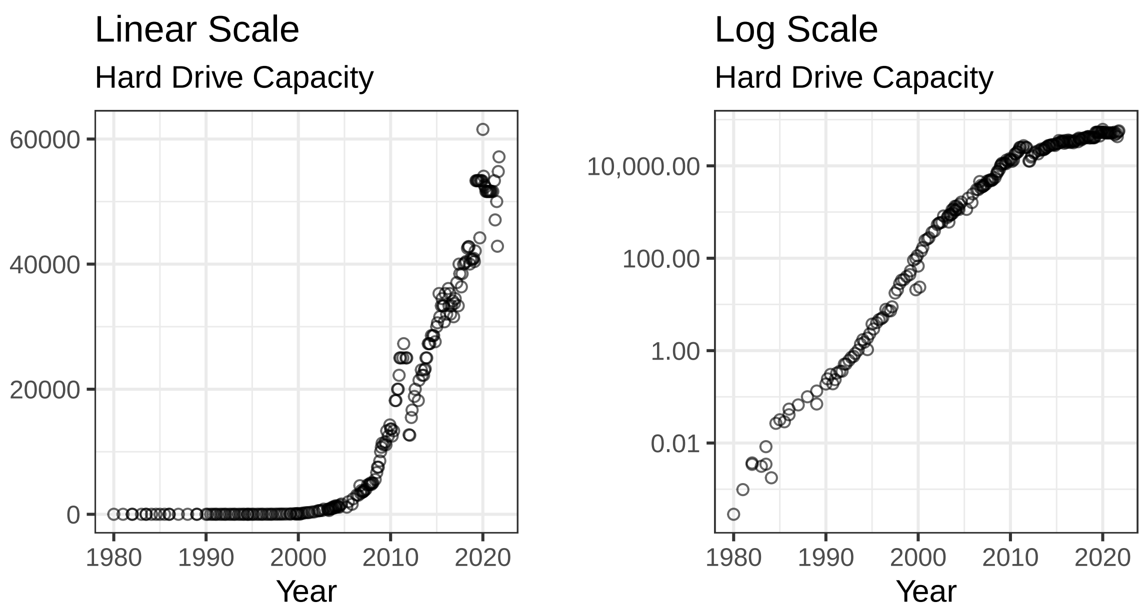

Figure 1: These plots present hard drive capacity over the past forty years on both the linear and log scale and illustrate the use of the log scale when displaying data which spans several magnitudes.

identify differences between exponentially increasing curves with varying levels of curvature using both linear and log scales. The study did not require participants receive any mathematical training or have prior understanding of exponential growth or logarithmic scales, focusing instead on the fundamental nature of the ability to identify differences in charts. In Section 2 we describe our participant sample, the graphical task to be completed, and the data generation process and study design. Section 3 describes the participant data collected and shares results from the statistical analyses of the data using a generalized linear mixed model. We present overall conclusions and discussion of results in Section 4, and provide an overview of future related papers. The Supplementary Material contains a link to the RShiny data collection applet, participant data used for analysis, and code to replicate the analysis. This results from this study lay the groundwork for further exploration of the implications of using log scales in data visualization.

1.1 Statistical Lineups

Buja et al. (2009) introduced statistical lineups as a framework for statistical inference and graphical tests. Statistical lineups treat a data plot as a visual statistic, summarizing the data as a numerical function or mapping. Evaluation of a panel in a statistical lineup requires visual inspection by a person, and if visual evaluations lead to different results, two visualization methods are deemed significantly different. Recent studies have utilized statistical lineups to quantify the perception of graphical design choices (Hofmann et al. 2012, Loy et al. 2017, 2016, VanderPlas & Hofmann 2017). Statistical lineups provide an elegant way of combining perception and statistical hypothesis testing through graphical experiments (Majumder et al. 2013, Vanderplas et al. 2020, Wickham et al. 2010).

The term 'lineup' is an analogy to police lineups in criminal investigations, where witnesses identify the criminal from a group of individuals. Similarly, researchers present a statistical lineup plot consisting of smaller panels and ask the viewer to identify the panel that contains the actual data from a set of decoy null plots. Researchers generate null plots containing data generated according to a prespecified hypothesis using permutation or simulation. Typically, a statistical lineup consists of 20 panels, with one target panel and 19 null panels. If the viewer can identify the target panel from the null panels, it

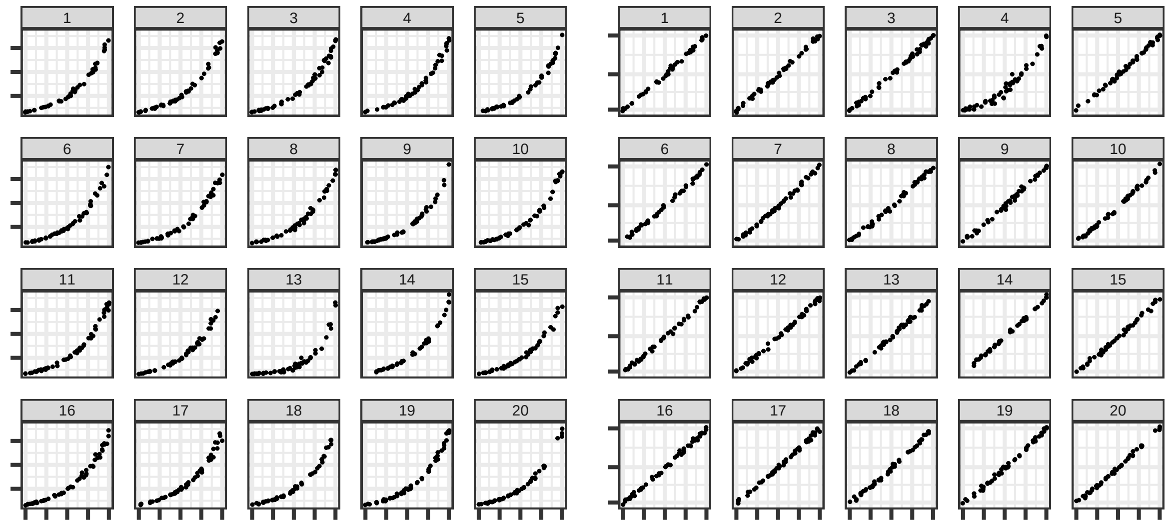

Figure 2: The lineup plot on the left displays increasing exponential data on a linear scale with panel (2 x 5) + 3 as the target. The lineup plot on the right displays increasing exponential data on the log scale with panel 2 x 2 as the target.

suggests that the actual data is visually distinct from the data generated under the null model. Fig. 2 presents examples of statistical lineups. The statistical lineup on the left presents increasing exponential data displayed on a linear scale with panel $(5 \times 2) + 3$ as the target. The lineup on the right shows increasing exponential data plotted on a log base ten scale with panel 2×2 as the target.

While explicit graphical tests direct the participant to a specific feature of a plot to answer a particular question, implicit graphical tests require the user to identify both the purpose and function of the plot in order to evaluate the plots shown. Furthermore, implicit graphical tests, such as lineups, simultaneously test for multiple visual features, including outliers, clusters, and linear and nonlinear relationships (VanderPlas & Hofmann 2015). Researchers can collect responses from multiple viewers using crowd-sourcing websites such as Prolific and Amazon Mechanical Turk.

2 Study Development and Methods

2.1 Data Generation

In this study, we simulated data from an exponential model to generate the target and null data sets; the models between panels differ in the parameter values selected for the null and target panels. In order to guarantee the simulated data spans the same domain and range of values for each statistical lineup panel, we began with 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. We mapped the randomly generated x values to a corresponding y value based on an exponential model with predetermined parameter values and multiplicative random errors to simulate the response. These constraints assure that participants who select the target panel 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:

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

The parameters α and θ were 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 in Algorithm 1 and Algorithm 2.

2.2 Parameter Selection

We followed a 'Goldilocks' inspired procedure to choose three levels of trend curvature (low curvature, medium curvature, and high curvature). For each curvature level, we simulated 1,000 data sets of (x_{ij}, y_{ij}) points for $i = 1, ..., 50$ increments of x-values and replicate $j = 1, ..., 10$ corresponding y-values per x-value. Each generated x_i point from Algorithm 2 was replicated ten times. We fit a linear regression model on each of the individual data sets and computed the lack of fit statistic (LOF) which measures the deviation of the

Algorithm 1 Lineup Parameter Estimation

- **Input Parameters:** domain $x \in [0, 20]$, range $y \in [10, 100]$, midpoint x_{mid}.
- **Output Parameters:** 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 of points (x_k, y_k) for $k = 1, 2, 3, 4$, calculate the coefficients from the linear regression 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 base `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 get estimated parameter values - $\hat{\alpha}, \hat{\beta}, \hat{\theta}$.

Algorithm 2 Lineup Exponential Data Simulation

- **Input Parameters:** sample size $N = 50$, estimated parameters $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\theta}$, from Algorithm 1, and standard deviation σ from the exponential curve.
- **Output Parameters:** N points, in the form of vectors $\mathbf{x}$ and $\mathbf{y}$.

1: Generate $\tilde{x}_j, j = 1, ..., \frac{3}{4}N$ 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 guarantees some variability and potential clustering 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 deviation 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 $e_i \sim N(0, \sigma^2)$.

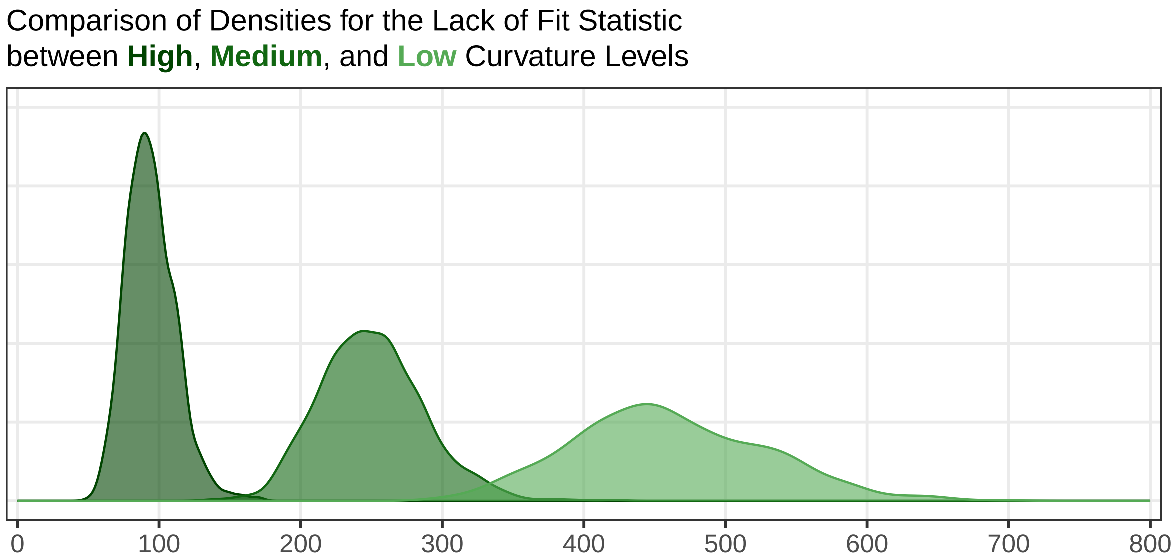

Figure 3: Density plot of the lack of fit statistic showing separation of difficulty levels: obvious curvature, noticable curvature, and almost linear.

Table 1: Lineup data simulation final parameters

Curvature Level	x_{mid}	$\hat{\alpha}$	$\tilde{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\sigma}$
High	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
Low	11.5	37.26	37.22	0.06	-27.26	0.05

data from the linear regression model. After obtaining the LOF statistic for each level of curvature, we evaluated the density plots (Fig. 3) to provide a metric for differentiating between the curvature levels and thus detecting the target plot. While the LOF statistic provides a numerical value for discriminating between the difficulty levels, it cannot be directly related to the perceptual discriminability; it serves primarily as an approximation to ensure that we are testing parameters at several distinct curvature levels. Table 1 lists the final parameters used for data simulation.

2.3 Lineup Setup

To generate the small multiple scatter plots for the statistical lineups shown to participants in the study, we simulated a single data set corresponding to curvature level A for the target plot and multiple data sets corresponding to curvature level B for the null plots. The `nullabor` package in R (Buja et al. 2009) randomly assigned the target plot to one of the panels surrounded by panels containing null plots. For example, the statistical lineup randomly embeds a target plot with simulated data following an increasing exponential curve with high curvature within null plots with simulated data following an increasing exponential trend with low curvature. The target and null panels spanned a similar domain and range due to the implemented constraints when simulating the data. There were a total of six lineup curvature combinations; Fig. 4 illustrates the six lineup curvature combinations (top: linear scale; bottom: log scale) where the green line indicates the curvature level designated to the target plot while the black line indicates the curvature level assigned to the null plots. Two sets of each lineup curvature combination were simulated (a total of twelve test data sets) and plotted on both the linear scale and the log scale (24 test lineup plots). In addition, three curvature combinations generated homogeneous "Rorschach" lineups, where all panels were from the same distribution. Each participant evaluated one "Rorschach" lineup, but for simplicity, we did not describe these evaluations, and we left their analysis to a later date.

2.4 Study Design

We showed each participant thirteen lineup plots (twelve test and one Rorschach). At the start of the study, we randomly assigned participants one of the two replicate data sets for each of the six unique lineup curvature combinations. Participants evaluated the lineup plot corresponding to the linear and log scales for each assigned test data set. For the additional Rorschach lineup plot, we randomly assigned participants one data set shown on either the linear or the log scale. We randomized the order in which participants saw the assigned thirteen lineup plots.

We used Prolific, a survey site that connects researchers to study participants, to recruit participants above the age of majority in their region. Following the comple-

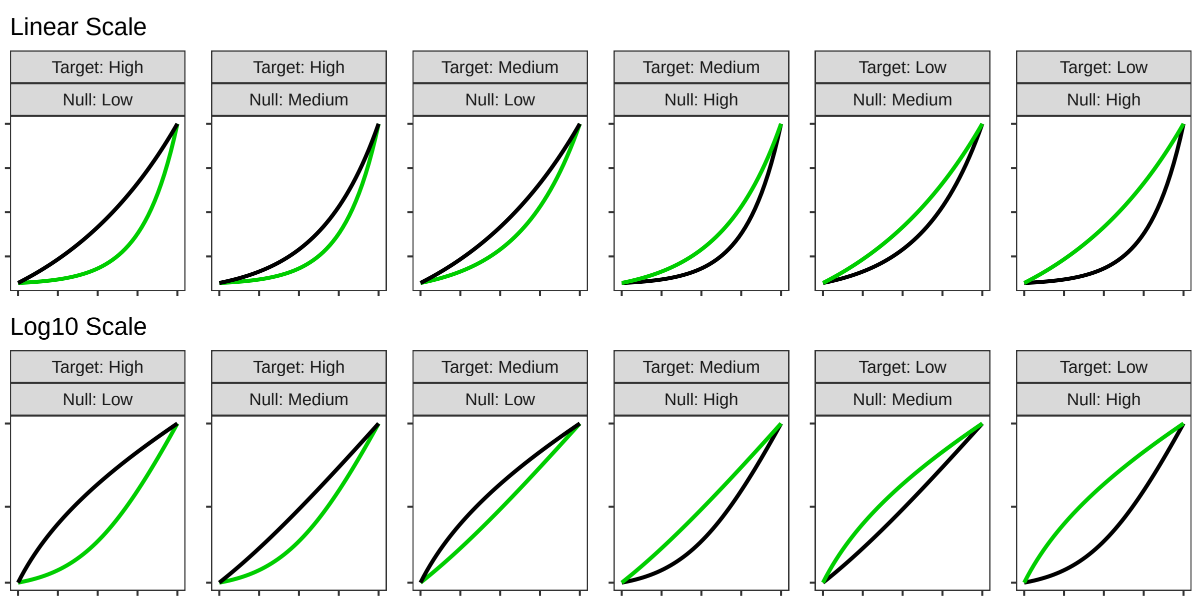

Figure 4: Thumbnail plots illustrating the six curvature combinations displayed on both scales (linear and log). The green line indicates the curvature level to be identified as the target plot from amongst a set of null plots with the curvature level indicated by the black line.

tion of the current study, participants sequentially completed two additional graphical experiments related to the perception of logarithmic perception (to be discussed at a later date), and we compensated them for their participation in the series of all three studies. While we did not request a representative sample, previous literature suggests that prior mathematical knowledge or experience with exponential data is not associated with the outcome of graphical experiments involving lineups (VanderPlas & Hofmann 2015). Participants completed the series of graphical tests using an R Shiny (Chang et al. 2022) application found at `https://shiny.srvanderplas.com/perception-of-statistical-graphics/`; the code used to create the study application can be accessed at `https://github.com/earobinson95/log-perception-prolific`.

The applet web page guided participants through a series of lineup plots and asked them to identify the plot which appeared to be most different from the others. In each lineup evaluation, participants justified their choice and provided their level of confidence in their choice. This graphical task aimed to test an individual's ability to perceptually differentiate exponentially increasing trends with differing levels of curvature on both the linear and log scale.

2.5 Statistical Analysis

Each lineup plot evaluated was assigned a binary value based on the participant response (correct target plot identification = 1, not correct target plot identification = 0). We defined Y_{ijkl} as the event that participant $l = 1, ..., N_{\text{participant}}$ correctly identified the target plot for data set $k = 1, 2$ with curvature combination $j = 1, 2, 3, 4, 5, 6$ plotted on scale $i = 1, 2$. The binary response was analyzed using a generalized linear mixed model (GLMM) following a binomial distribution with a logit link function with a row-column blocking design accounting for the variation due to participant and data set, respectively, as

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

where

- η is the baseline average probability of selecting the target plot
- δ_i is the effect of scale $i = 1, 2$

- γ_j is the effect of curvature combination $j = 1, 2, 3, 4, 5, 6$
- $\delta\gamma_{ij}$ is the two-way interaction between the i^{th} scale and j^{th} curvature combination
- $s_l \sim N(0, \sigma^2_{\text{participant}})$ is the random effect for participant characteristics
- $d_k \sim N(0, \sigma^2_{\text{data}})$ is the random effect for data specific characteristics.

We assumed that random effects for data set and participant are independent. Target plot identification was analyzed using a GLMM implemented in `glmer` from the `lme4` R (version 4.2.2) package (Bates et al. 2015). We used the `emmeans` R package (Lenth 2021) to calculate the estimated target detection probabilities and obtain odds ratio comparisons between the log and linear scale.

3 Results

Supplement participant demographics (broad location info, etc)

We recruited participants and conducted the study via Prolific, a crowd-sourcing website, in March 2022, during which 325 individuals completed 4,492 individual test lineup evaluations. We included only participants who completed the entire study in the final analysis, which included 311 participants and 3,958 lineup evaluations; due to server capacity, some participants were required to restart the study, thus resulting in the possibility of more than twelve lineup evaluations per participant. As a whole, participants evaluated each uniquely generated lineup plot between 141 and 203 times (Mean: 164.92, SD: 14.9). Participants correctly identified the target panel in 47% of the 1,981 lineup evaluations made on the linear scale and 65.3% of the 1,977 lineup evaluations made on the log scale.

The results from the GLMM indicated a strong interaction between the curvature combination and scale ($\chi^2_5 = 294.443$; $p < 0.0001$), and estimated variance due to participant and data set to be $\hat{\sigma}^2_{\text{participant}} = 1.19$ (s.e. $= 1.09$) and $\hat{\sigma}^2_{\text{data}} = 0.433$ (s.e. $= 0.66$), respectively. Therefore, we concluded there was low variability in the accuracy of target panel detection between participants and across replications of uniquely simulated data sets. To determine the effect of scale, we compared the estimated accuracy between the log and linear scale within each curvature combination due to the interaction as determined by the GLMM.

On both the log and linear scales, the highest accuracy occurred in lineup plots where the target model and null model had a considerable difference in curvature, and the target plot had more curvature than the null plots (high curvature target plot embedded in low curvature null plots). There is a decrease in accuracy on the linear scale when comparing a target plot with less curvature to null plots with more curvature (medium curvature target plot embedded in high curvature null plots; low curvature target plot embedded in medium curvature null plots; low curvature target plot embedded in high curvature null plots). Best et al. (2007) found that the accuracy of identifying the correct curve type was higher when presented with nonlinear trends, 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. Fig. 5 displays the estimated (log) odds ratio of successfully identifying the target panel on the log scale compared to the linear scale. The thumbnail figures to the right of the plot illustrate the curvature combination on the linear (left thumbnail) and log base ten (right thumbnail) scales associated with the y-axis label. The choice of scale had no impact if there curvature differences were substantial and the target plot had more curvature than the null plots (high curvature target plot embedded in low curvature null plots). However, presenting data on the log scale makes us more sensitive to slight changes in curvature (low or high curvature target plot embedded in medium curvature null plots; medium curvature target plot embedded in high curvature null plots) and apparent differences in curvature when the target plot had less curvature than the null plots (low curvature target plot embedded in high curvature null plots). An exception occurred when identifying a plot with curvature embedded in null plots close to a linear trend (medium curvature target panel embedded in low curvature null panels). The results indicate that participants were more accurate at detecting the target panel on the linear scale than on the log scale. When examining this curvature combination, the same perceptual effect occurred as we previously saw, but in a different context of scales. On the linear scale, participants were perceptually identifying a convexly curved trend from close to a linear trend, whereas after the logarithmic transformation, participants were perceptually identifying a trend

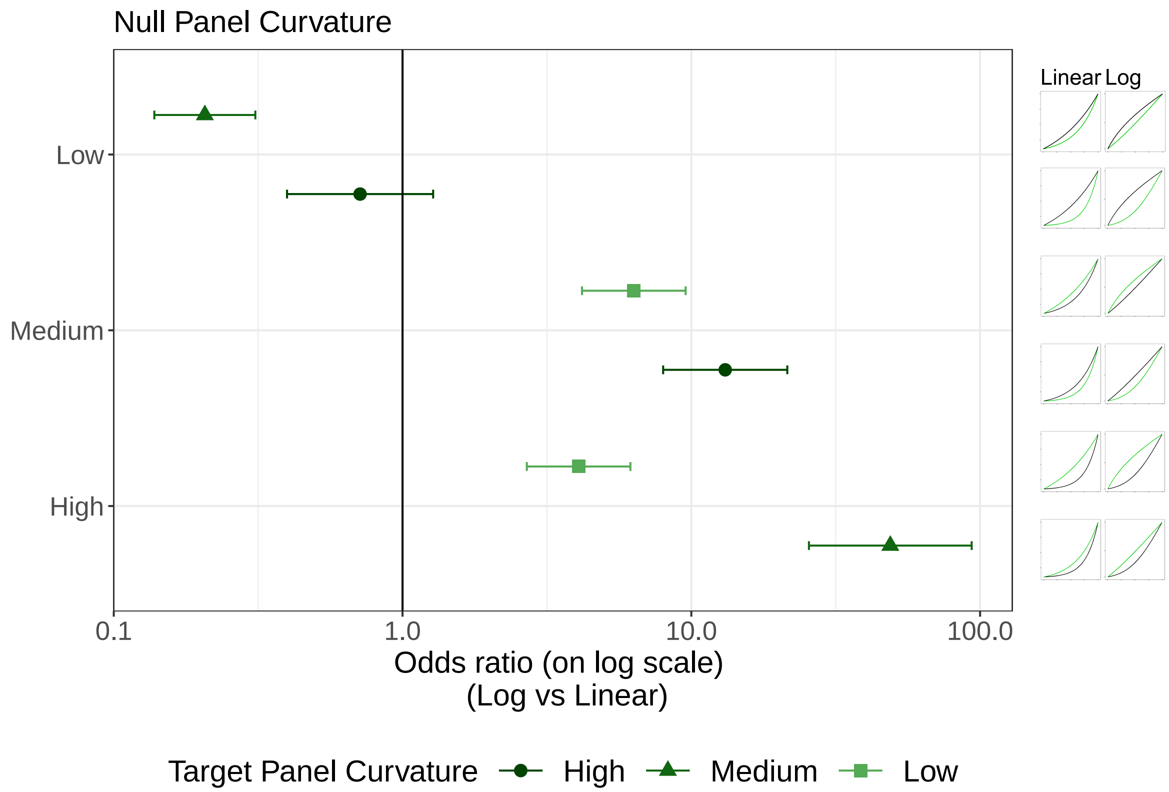

Figure 5: Estimated (log) odds ratio of successfully identifying the target panel on the log scale compared to the linear scale. The y-axis indicates the the model parameters used to simulate the null plots with the target plot model parameter selection designated by shape and shade of green. The thumbnail figures on the right display the curvature combination as shown in Fig. 4 on both scales (linear - left, log - right).

close to linear from a concavely curved trend Fig. 4. This result again supports the claim that it is easy to identify a curve in a bunch of lines but harder to identify a line in a bunch of curves (Best et al. 2007).

4 Discussion and Conclusion

This work aims to provide foundational research to support the principles that guide design decisions in scientific visualizations of exponential data. In this study, we explored the use of linear and log scales to determine whether the choice of scale impacts our ability to notice differences in exponentially increasing trends. The results indicated that when there was a considerable difference in curvature between the target plot and null plots and the target plot had more curvature than the 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 or apparent curvature differences when the target plot had less curvature than the null plots. 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. Using visual inference to identify these guidelines suggests that there are perceptual advantages to log scales when differences are subtle.

We conducted this study as the first in a series of three graphical tests to understand the perceptual and cognitive implications of using log scales to display exponentially increasing data. In our next two papers in this series, we will investigate whether using log scales presents cognitive disadvantages, such as making it harder to utilize graphical information. These studies serve as an example of multi-modal graphical testing, examining different levels of engagement and interaction with graphics in order to produce nuanced, specific guidelines for graphical design. By testing graphics in situations similar to how they are used, we can gain additional insight into graphical perception and improve visual communication of scientific results.

Supplementary Material

- **Participant Data:** De-identified participant data collected in the study and used for analyses (lineup-model-data.csv).
- **Data Analysis Code:** The code used to replicate the analysis in this paper (lineups-analysis.qmd)
- **README:** File containing detailed descriptions of the supplementary material (README.html).

References

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015), 'Fitting linear mixed-effects models using lme4', *Journal of Statistical Software* **67**(1), 1–48.

Best, L., Smith, L. & Stubbs, D. (2007), 'Perception of linear and nonlinear trends: Using slope and curvature information to make trend discriminations', *Perceptual and Motor Skills* **104**(3), 707–721.

Buja, A., Cook, D., Hofmann, H., Lawrence, M., Lee, E.-K., Swayne, D. & Wickham, H. (2009), 'Statistical inference for exploratory data analysis and model diagnostics', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **367**(1906), 4361–4383.

Burn-Murdoch, J., Nevitt, C., Tilford, C., Rininsland, A., Kao, J., Elliott, O., Lewis, E., Fox, B. & Stabe, M. (2020), 'Coronavirus tracked: Has the epidemic peaked near you?'.
URL: *https://ig.ft.com/coronavirus-chart/?areas=eur*

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A. & Borges, B. (2022), *shiny: Web Application Framework for R*. R package version 1.7.4.
URL: *https://CRAN.R-project.org/package=shiny*

Fagen-Ulmschneider, W. (2020), '91-divoc'.
URL: *https://91-divoc.com/pages/covid-visualization/*

Heckler, A., 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)', IEEE, pp. 1066–1071.

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.

Lenth, R. (2021), *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.6.0.
URL: *https://CRAN.R-project.org/package=emmeans*

Loy, A., Follett, L. & Hofmann, H. (2016), 'Variations of q–q plots: The power of our eyes!', *The American Statistician* **70**(2), 202–214.

Loy, A., Hofmann, H. & Cook, D. (2017), 'Model choice and diagnostics for linear mixed-effects models using statistics on street corners', *Journal of Computational and Graphical Statistics* **26**(3), 478–492.

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.

Menge, D., MacPherson, A., Bytnerowicz, T., Quebbeman, A., Schwartz, N., Taylor, B. & Wolf, A. (2018), 'Logarithmic scales in ecological data presentation may cause misinterpretation', *Nature Ecology & Evolution* **2**(9), 1393–1402.

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. (2015), 'Spatial reasoning and data displays', *IEEE Transactions on Visualization and Computer Graphics* **22**(1), 459–468.

VanderPlas, S. & Hofmann, H. (2017), 'Clusters beat trend!? testing feature hierarchy in statistical graphics', *Journal of Computational and Graphical Statistics* **26**(2), 231–242.

Waddell, W. (2005), 'Comparisons of thresholds for carcinogenicity on linear and logarithmic dosage scales', *Human & Experimental Toxicology* **24**(6), 325–332.

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