Perception of Curvature & Exponential Growth

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

September 20, 2022

Abstract

The text of your abstract. 200 or fewer words.

Keywords: log scales, visual inference, graphical testing

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. Identifying differences 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. The study in this chapter is conducted through visual inference and the use of statistical lineups (Buja et al. 2009*a*) to differentiate between exponentially increasing curves with differing levels of curvature, using linear and log scales.

2 Visual Inference

In Section 1.4, we explained how a data plot can be evaluated and treated as a visual statistic, a numerical function which summarizes the data. To evaluate a graph, the statistic (data plot) must be run through a visual evaluation - a person. We can conclude that two visualization methods are significantly different if the visual evaluation is different. Recent graphical experiments 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 using graphical experiments (Majumder et al. 2013, Vanderplas et al. 2020, Wickham et al. 2010). '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 where the viewer is asked to identify the panel containing the real data from among a set of decoy null plots. Null plots display data under the assumption there is no relationship and can be generated by permutation or simulation. A statistical lineup typically consists of 20 panels - one target panel and 19 null panels. If the viewer can identify the target panel randomly embedded within the set of null panels, this suggests that the real data is visually distinct from data generated under the null model. Fig. 1 provides examples of statistical lineups. The lineup plot on the left displays increasing exponential data displayed on a linear scale with panel 13 as the target; the lineup plot on the right displays increasing exponential data on the

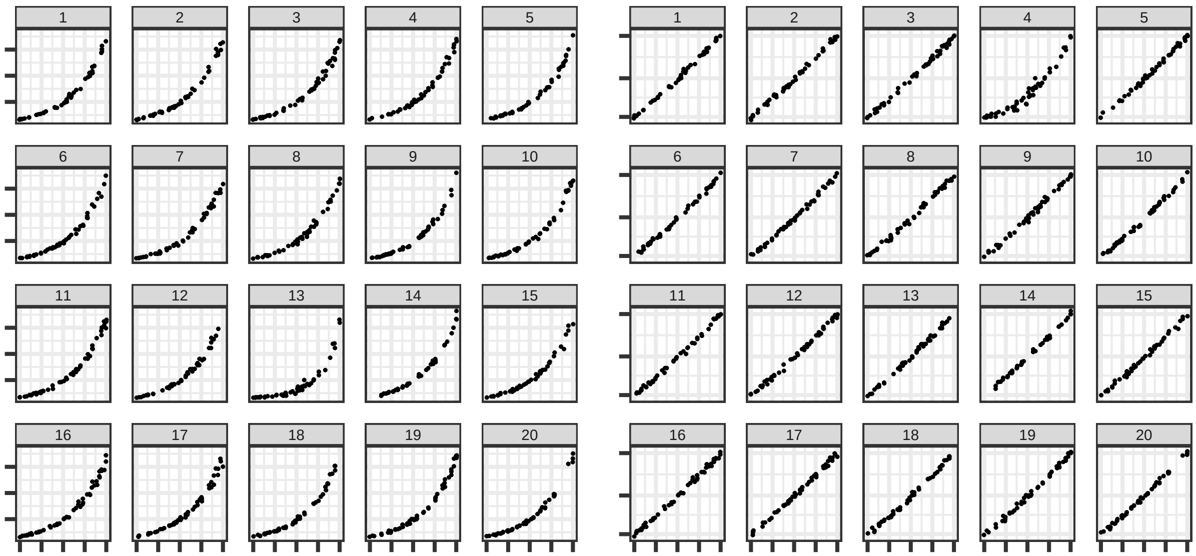

Figure 1: The lineup plot on the left displays increasing exponential data on a linear scale with panel (2 x 5) + 3 as the target. The lineup plot on the right displays increasing exponential data on the log scale with panel 2 x 2 as the target.

log base ten scale with panel 4 as the target.

While explicit graphical tests direct the participant to a specific feature of a plot to answer a specific question, implicit graphical tests require the user to identify both the purpose and function of the plot in order to evaluate the plots shown (Vanderplas et al. 2020). Implicit graphical tests, such as lineups, have the advantage of simultaneously visually testing for multiple visual features including outliers, clusters, linear and nonlinear relationships. Responses from multiple viewers are collected through convenience sampling (in informal situations) or crowd sourcing websites such as Prolific, Amazon Mechanical Turk, and Reddit (in more formal situations).

3 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 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 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.

Data were simulated 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.

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 - $\hat{\alpha}, \hat{\beta}, \hat{\theta}$.

4 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

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

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. On each of the individual data sets, we fit a linear regression model and computed the lack of fit statistic (LOF) which measures the deviation of the data from the linear regression model. The density curves of the LOF statistics for each level of curvature are plotted (Fig. 2) 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. Final parameters used for data simulation are shown in Table 1.

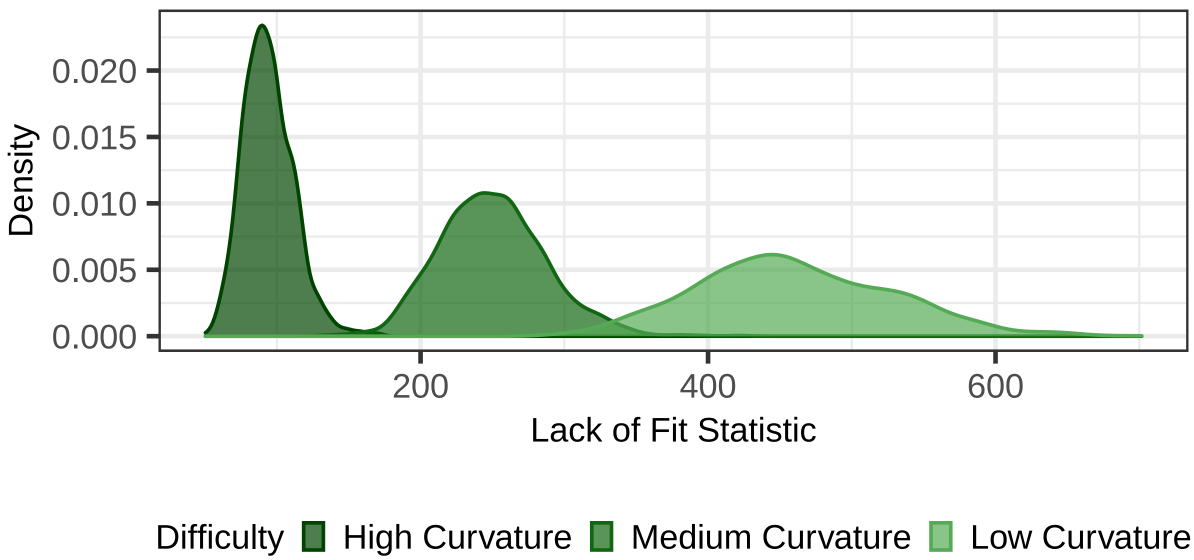

Figure 2: Density plot of the lack of fit statistic showing separation of difficulty levels: obvious curvature, noticable curvature, and almost linear.

Table 1: Lineup data simulation final parameters

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

5 Lineup Setup

Lineup plots were generated by mapping one simulated data set corresponding to curvature level A to a scatter plot to be identified as the target panel while multiple simulated data sets corresponding to curvature level B were individually mapped to scatter plots for the null panels. The `nullabor` package in R (Buja et al. 2009*b*) was used to randomly assign the target plot to one of the panels surrounded by panels containing null plots. For example, a target plot with simulated data following an increasing exponential curve with high curvature is randomly embedded within null plots with simulated data following an increasing exponential trend with low curvature. By the implemented constraints, the target panel and null panels spanned a similar domain and range. There were a total of six lineup curvature combinations; Fig. 3 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 (total of twelve test data sets) and plotted on both the linear scale and the log scale (total of 24 test lineup plots). In addition, there were three curvature combinations which generated homogeneous "Rorschach" lineups, where all panels were from the same distribution. Each participant evaluated one of these lineups, but for simplicity, these evaluations are not described in this chapter and their analysis is left to a later date.

6 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 curvature 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 in their region were recruited from Prolific, a

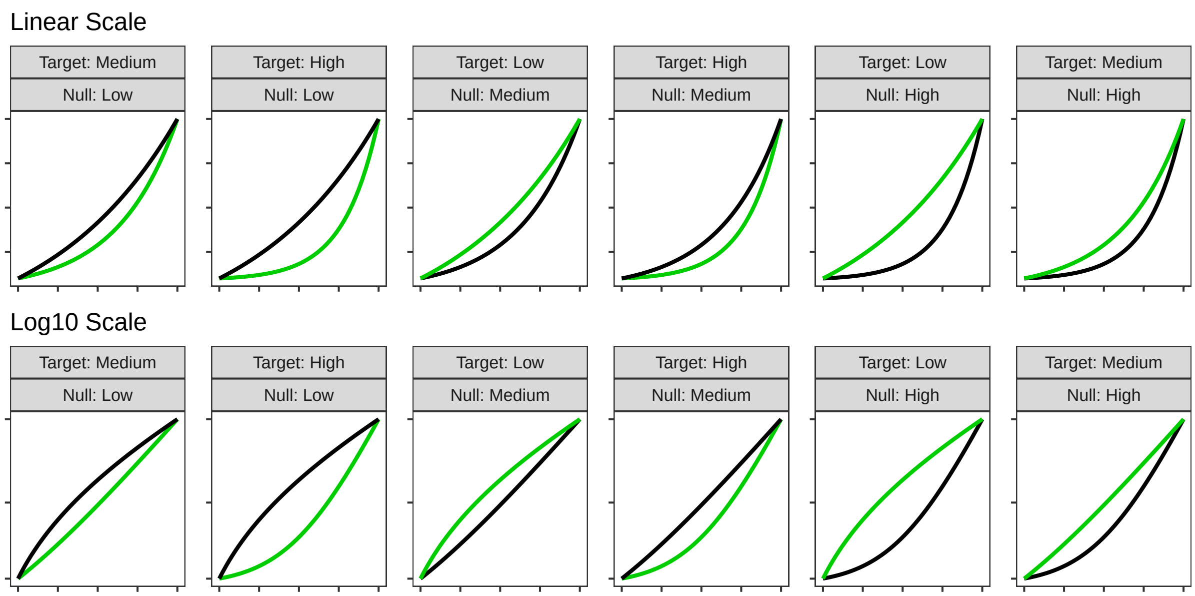

Figure 3: Thumbnail plots illustrating the six curvature combinations displayed on both scales (linear and log). The green line indicates the curvature level to be identified as the target plot from amongst a set of null plots with the curvature level indicated by the black line.

survey site that connects researchers to study participants. Participants were compensated for their time and participated in all three related graphical studies consecutively. 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). The lineup study in this chapter was completed first in the series of graphical studies.

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 graphical task was to test an individual's ability to perceptually differentiate exponentially increasing trends with differing levels of curvature on both the linear and log scale.

7 Results

Participant recruitment and study deployment were conducted via Prolific, a crowd sourcing website, on Wednesday, March 23, 2022 during which 325 individuals completed 4,492 unique test lineup evaluations. Only participants who completed the lineup study were included in the final data set which included a total of 311 participants and 3,958 lineup evaluations. Each plot was evaluated 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.

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} to be 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 package (Bates et al. 2015). Estimates and odds ratio comparisons between the log and linear scales were calculated using the `emmeans` R package (Lenth 2021).

Results indicated a strong interaction between the curvature combination and scale ($\chi^2_5 = 294.443$; p < 0.0001). Variance due to participant and data set were estimated 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.

On both the log and linear scales, the highest accuracy occurred in lineup plots where the target model and null model had a large curvature difference 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 accuracy of identifying the correct curve type was higher when nonlinear trends were presented indicating that it is hard to say something is linear (something has less curvature), but easy to say that it is not linear; our results concur with this observation. Fig. 4 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 both 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 curvature differences are large 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 large 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 occured 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 the log scale. When examining this curvature combination, the same perceptual effect occurred as what we previously saw, but in a different context of scales. On the linear scale, participants were perceptually identifying a curved trend from close to a linear trend whereas after the logarithmic transformation, participants were perceptually identifying a trend close to linear from a curved trend. This 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).

8 Discussion and Conclusion

The overall goal of this chapter 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 exponentially increasing trends is impacted by the choice of scale. The results indicated that when there was a large 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 large 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. The use of visual inference to

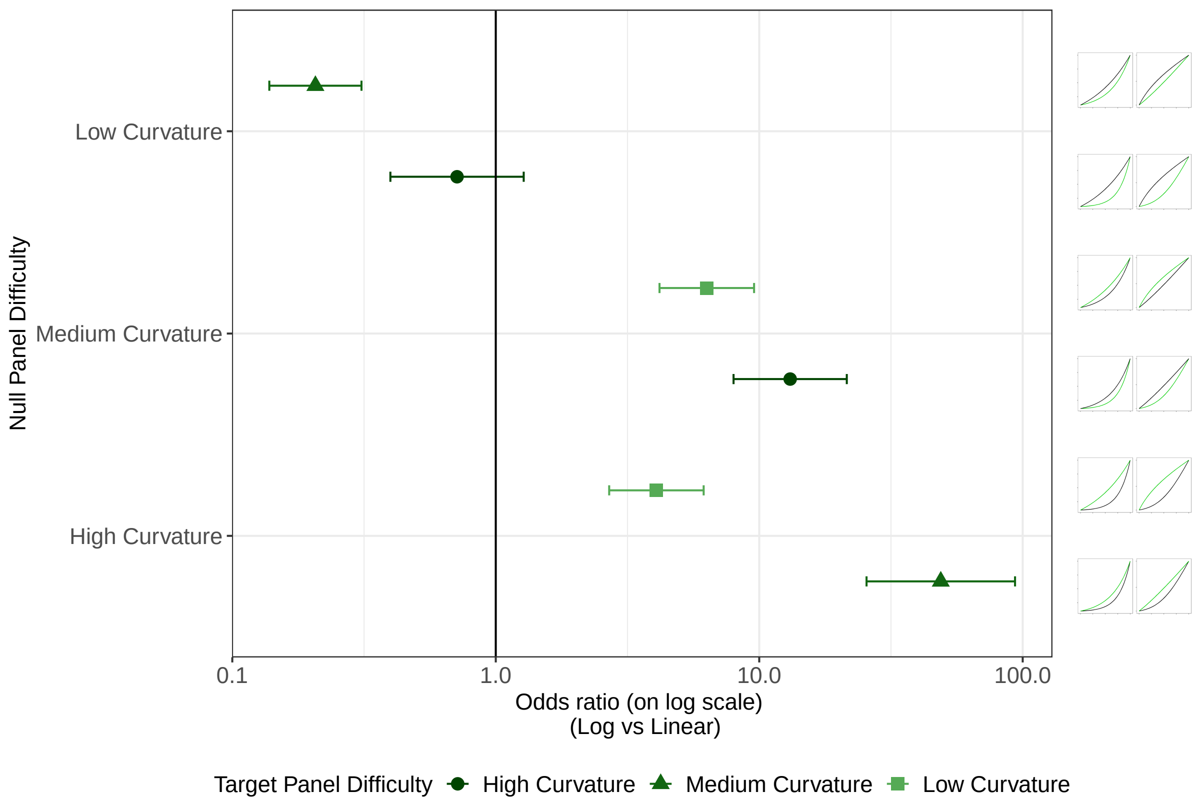

Figure 4: Estimated (log) odds ratio of successfully identifying the target panel on the log scale compared to the linear scale. The y-axis indicates the the model parameters used to simulate the null plots with the target plot model parameter selection designated by shape and shade of green. The thumbnail figures on the right display the curvature combination as shown in Fig. 3 on both scales (linear - left, log - right).

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?

References

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015), 'Fitting Linear Mixed-Effects Models Using lme4', *Journal of Statistical Software* **67**, 1–48.
URL: *https://doi.org/10.18637/jss.v067.i01*

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*a*), '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.

Buja, A., Cook, D., Hofmann, H., Lawrence, M., Lee, E.-K., Swayne, D. & Wickham, H. (2009*b*), '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.

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.

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.

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