Project Description

1 Objectives

Statistical graphics are powerful and efficient tools to convey data to support human decision making; however, empirical research on graphical perception is sparse relative to the number of decisions necessary to make a good chart; when relevant studies are available, they often use incomparable methods and produce conflicting results. Many recommendations are based on opinion rather than empirical study, rendering scientific communications sub-optimal or ineffective. This is alarming, as effective science communication is critical for cultivating public trust in the scientific process and ensuring that decision makers accurately interpret information when making choices which impact people's lives.

Our visual systems are uniquely optimized for efficient assessment of visual information, which is why data visualizations are so effective. Some disciplines, such as forensic pattern analysis (shoe-prints, firearms, handwriting) take this further, using subjective visual evaluations in place of quantitative methods. Fundamentally, these comparisons are similar to data visualization: the image encodes numerical data and actionable information is extracted via human perception. We can use the insights we gain by studying perception of graphics to "close the loop", building empirical, explainable features for pattern evidence that replicate visual comparisons performed by forensic examiners, producing machine-learning algorithms to provide objective, quantitative support for examiner testimony in court.

This CAREER proposal addresses a fundamental research question underpinning both of these problems: *How do humans make decisions supported by data visualizations?* Three research objectives (ROs) will support the overall **research goal** of advancing our understanding of the role of data visualizations in human decision making:

- **RO1:** Examine effectiveness of charts across different tasks, such as visual search, comparison, prediction, and estimation, by developing multi-modal methods for graphical testing. If successful, this will produce testing methods that focus on specific cognitive tasks related to use of scientific visualizations, and will facilitate comparison of results across experiments and methods.
- **RO2:** Empirically validate visualization design guidelines, measuring the impact of design decisions on performance of different tasks. If successful, this will produce nuanced guidelines that are task-focused and empirically derived.
- **RO3:** Engineer statistical features that mimic human perception which will be combined into empirically validated, objective, quantitative tools that support reproducible decision-making in forensic pattern evidence.

Integrated with these research efforts, the overall **education goal** is to cultivate statistical learning and scientific decision making in society. Three education objectives (EOs) address this goal:

- **EO1:** Develop and implement experiential learning activities in graphics for undergraduate introductory statistics courses.
- **EO2:** Create and implement open educational resources (OER) to introduce reproducible science and open-source software development in statistical programming courses.
- **EO3:** Engage with forensic scientists, lawyers, and judges, evaluating the scientific support for forensic disciplines and promoting open, reproducible scientific practices.

The research, educational activities, and outreach conducted in support of these objectives will advance our understanding of the design, perception, and use of visual data displays. New suites of task-focused testing methods will facilitate precise, targeted experiments in data visualization, and applying multiple task-focused testing methods to examine design guidelines will facilitate better scientific communication and dissemination of results, building trust in science and leveraging scientific principles. Measuring the pedagogical effectiveness of experiential learning in undergraduate statistics courses will provide new understanding of effective strategies for STEM outreach and engagement. Incorporating students into the research process with experiential learning and using graphics to motivate statistical concepts will engage students at a deeper level and open up STEM topics to math-phobic students in a tangible way. Weaving scientific communication and reproducibility into courses on statistical programming will produce a new generation of scientists equipped with the tools to conduct better, more reproducible research and communicate the results to other scientists and to the public.

2 Research Plan

2.1 Background Information

Statistical visualizations are a primary way scientists communicate results to other scientists as well as the general public, but we have been arguing about best practices for the visual display of data almost since these graphical forms were created [1–3]. Different individuals and organizations have released guidelines for graphical design [4–8], but many of these are based on aesthetic preferences, individual experience, or extrapolation from limited empirical studies. Even when there are empirical studies to support guidelines for choosing one graphical form over another [9–11, see 12 for a review], these studies measure different quantities in different ways ; the design and implementation differences have a potentially large effect on the generalizability of the results. Complicating this problem, the design space of visualization user studies is incredibly large [13], and studies can address charts from multiple perspectives, including perception, task-based interaction (with different tasks producing different results), structural characteristics, and more [14]. In addition, the test population is an important and under-examined aspect of the graphics experiment design space: we know that expertise as well as disorders such as dyslexia, dyscalculia, and ADHD affect perception, numeracy, and other processes involved in graph comprehension [15–17]; while these subpopulations can be difficult to test compared to the general population, empirical design recommendations must also address accessibility and audience considerations, as designers already consider these issues [18] but have no firm guidance about best practice. Relatively few studies directly compare the effect of different approaches for testing graphics, but minor changes to data collection methods can impact accuracy of even simple numerical estimation tasks [19]; it is reasonable to expect that the specifics of the graphical testing procedure used may impact the results, potentially explaining conflicting empirically based recommendations [2, 3, 11]. Without the ability to integrate results from different types of studies, it is hard to generate scientifically accurate recommendations from existing research. To bridge this gap, information visualization needs a suite of easily implemented, validated graphical testing methods. These methods should then be systematically applied to test existing graphical design guidelines, with the intent to produce nuanced recommendations backed by tailored empirical research.

2.2 Preliminary Data

During the early stages of the pandemic, log scales were fairly common because they allow comparison of different orders of magnitude on the same chart; however, it was difficult to find studies evaluating how successfully log scales were read and interpreted, though there were studies suggesting that people systematically under-predict exponential growth [20, 21]. In order to address this gap in the literature, I worked with a graduate student, Emily Robinson, to design and execute a series of three studies addressing different

tasks related to log scales: could participants detect a difference in rates of exponential growth? Could they predict future values? Could they estimate additive and multiplicative differences from an exponential time series? We examined these questions by testing charts with log and linear scales; in addition to a separate study testing each question above [22, peer reviewed articles in progress], we have also published a tool paper [23] and a method validation paper [24].

2.3 Methods

3 Education Plan

3.1 Overview

3.2 Design and methods

3.3 Evaluation

3.4 Integration of research and Education

4 Broader Impacts
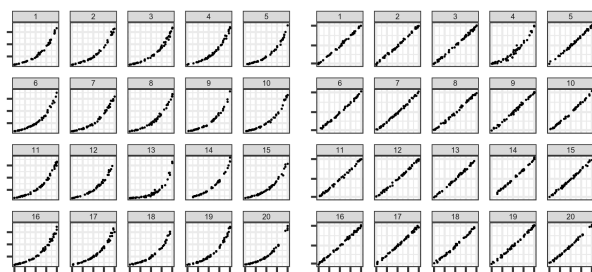
5 Timeline

6 Results from Prior NSF Support

Figure 1: Lineups (log and linear scale) test detection of differences in exponential growth rates.

References

1. Wickham H (2013) Graphical criticism: Some historical notes. *Journal of Computational and Graphical Statistics*, 22(1):38–44. https://doi.org/10.1080/10618600.2012.761140

2. Croxton FE, Stryker RE (1927) Bar Charts Versus Circle Diagrams. *Journal of the American Statistical Association*, 22(160):473–482. https://doi.org/10.2307/2276829

3. Eells WC (1926) The Relative Merits of Circles and Bars for Representing Component Parts. *Journal of the American Statistical Association*, 21(154):119–132. https://doi.org/10.2307/2277140

4. (1915) Joint committee on standards for graphic presentation. *Quarterly Publications of the American Statistical Association*, 14(112):790–797. https://doi.org/10.1080/15225445.1915.10503668

5. Brewer CA (1994) Guidelines for use of the perceptual dimensions of color for mapping and visualization. *Color hard copy and graphic arts III*, 2171:54–63. https://doi.org/10.1117/12.175328

6. Kosslyn SM (2006) Graph design for the eye and mind.

7. Tufte E (2001) The visual display of quantitative information.

8. Kelleher C, Wagener T (2011) Ten guidelines for effective data visualization in scientific publications. *Environmental Modelling & Software*, 26(6):822–827. https://doi.org/10.1016/j.envsoft.2010.12.006

9. Cleveland WS, McGill R (1984) Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):531–554. https://doi.org/10.1080/01621459.1984.10478080

10. 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. https://doi.org/10.1109/TVCG.2012.230

11. Spence I (1990) Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance*, 16:683–692. https://doi.org/10.1037/0096-1523.16.4.683

12. Vanderplas S, Cook D, Hofmann H (2020) Testing statistical charts: What makes a good graph? *Annual Review of Statistics and Its Application*, 7(1):61–88. https://doi.org/10.1146/annurev-statistics-031219-041252

13. Abdul-Rahman A, Chen M, Laidlaw DH (2020) A survey of variables used in empirical studies for visualization. *Foundations of data visualization*, :161–179. https://doi.org/10.1007/978-3-030-34444-3_7

14. Bolte F, Bruckner S (2020) Measures in visualization space. *Foundations of data visualization*, :39–59. https://doi.org/10.1007/978-3-030-34444-3_3

15. Cheng D, Xiao Q, Chen Q, Cui J, Zhou X (2018) Dyslexia and dyscalculia are characterized by common visual perception deficits. *Developmental Neuropsychology*, 43:497–507. https://doi.org/10.1080/87565641.2018.1481068

16. Chity N, Harvey J, Quadri S, Pete S, Stein S (2012) Thinking DIfferently. Assistive Technology as a Complement to the Learning Style of Post-Secondary Students with ADHD: Recommendations for Design Opportunities.

17. Hokken MJ, Krabbendam E, van der Zee YJ, Kooiker MJG (2023) Visual selective attention and visual search performance in children with CVI, ADHD, and Dyslexia: A scoping review. *Child Neuropsychology*, 29(3):357–390. https://doi.org/10.1080/09297049.2022.2057940

18. Bako HK, Liu X, Battle L, Liu Z (2023) Understanding how Designers Find and Use Data Visualization Examples. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):1048–1058. https://doi.org/10.1109/TVCG.2022.3209490

19. Honda H, Kagawa R, Shirasuna M (2022) On the round number bias and wisdom of crowds in different response formats for numerical estimation. *Scientific Reports*, 12(1):8167. https://doi.org/10.1038/s41598-022-11900-7

20. Timmers H, Wagenaar WA (1977) Inverse statistics and misperception of exponential growth. *Perception*

& Psychophysics, 21(6):558–562. https://doi.org/10.3758/bf03198737

21. Wagenaar WA, Sagaria SD (1975) Misperception of exponential growth. *Perception & Psychophysics*, 18(6):416–422. https://doi.org/10.3758/BF03204114

22. Robinson EA (2022) Human Perception of Exponentially Increasing Data Displayed on a Log Scale Evaluated Through Experimental Graphics Tasks.

23. Robinson EA, Howard R, VanderPlas S (2023) "You Draw It": Implementation of Visually Fitted Trends with R2d3. *Journal of Data Science*, 21(2):281–294. https://doi.org/10.6339/22-JDS1083

24. Robinson EA, Howard R, VanderPlas S (2022) Eye Fitting Straight Lines in the Modern Era. *Journal of Computational and Graphical Statistics*, 0(0):1–8. https://doi.org/10.1080/10618600.2022.2140668