Data Visualization Best Practice Literature Review
Applied Visualization Research to Unethical AI Reporting

Sommer Harris
harris.som@northeastern.edu

Nolen Bryant
bryant.n@northeastern.edu

Abstract

This paper reviews literature on best practices in data visualization of irresponsible Artificial Intelligence (AI) instances. The introduction speaks about data visualization's ability to communicate data as well as to aid in development of new data and insights. We explain our motivation in how the irresponsible Artificial Intelligence dataset can help domain experts support more accountability in AI. We explore identified risks in data visualization, and present our countermeasure through catering our tool to the nuances of our data and user stories. We outline our data and ideal user, review best practices, then outline proposed features that would enable our data visualization to be a successful tool.

1. Introduction

Data driven decision making is becoming the definitive, meritorious standard across disciplines and applications. With increasing prevalence, the data is gathered 'in the wild' or from real world sources as opposed to that from traditional prospective studies where the data has a narrow focus and design determined a priori for the purpose of answering specific hypothesis.

By contrast, modern data sets can be large and complex, often highly multi-dimensional. We wear smart clothing which records our heart rate and daily step count. Our computer human interactions can be recorded. Businesses invest millions in consumer background data. In a context where more data is always available, finding models to simplify and optimize information retrieval is paramount. Data visualization resources play a significant role in the process of information acquisition, as they provide a much shorter pathway for the content reading cognitive process [5].

Storytelling for data can help readers to understand high-complexity science based subjects; The United Nations, among others, recommended the use of data storytelling as the main approach to publishing complex data [22]. Many case studies have found this to be the most effective method for explaining findings with non-experts of a given topic [9], and there have been documented effects of data visualization usage having a significantly higher persuasive power [18]. Research in the field of data visualization is often framed in terms of how it helps to "reveal" knowledge , support narrative storytelling , or otherwise facilitate pathways to "insight" [15].

Yet visualization should not be limited in scope to just be an end product of scientific analysis. It is also an exploration tool that scientists can use throughout the research life cycle. Many researchers note that new database technologies, coupled with emerging Web-based technologies, may hold the key to lowering the cost of visualization generation and allow it to become a more integral part of the scientific process [10].

As with any tool or process, there must be consideration given both to the most effective means of usage as well as the potential problems, bias or misuses that may arise. This is the case with both the topic of our summer research, as it is with the subject of our data, unethical instances of AI.

1.1. Motivation

As an increasing amount of decisions and actions are being made by computers, there is a higher potential for the the wrong decisions to be made. When combined with the fact that these AI generated decisions are increasingly impactful on society, the potential harm done increases as well.

As AI systems are deployed by companies, nonprofits, and governments at an increasing rate, impacting millions- and perhaps billions- of people, the question of responsibility in AI creation and use is of utmost importance [19]. In today's AI systems we see many irresponsible results such as perpetuation of racial bias, dangerous self-driving cars, inappropriate content on children's services, and scheduling systems that ignore workers' needs [14]. These systems are not legal or moral agents and cannot be held accountable the way a creator or user can [19]. Raji et al define accountability as "the state of being responsible or answerable for a system, its behavior and its potential impacts". They then say that a specific structure is needed to effectively monitor the ethical compliance of AI systems [19]. Laato et al. mentions that one key research path in AI is to focus on the unwanted effects of AI [16].

Data visualization can be a way to support analysts in communicating cautions around AI use. With the irresponsible AI atlas, we aim to build a tool that supports the effort to evaluate and support accountability. Whether these audits of AI systems are coming from within the organization, academia, or designated accountability bodies, this tool is for providing structure to experts in AI ethics, as they back up their recommendations with storytelling through data. At this stage of irresponsible AI auditing, data visualization is an ideal tool because it can support the continued research that needs to be done by experts and will support storytelling to the broader community once conclusions are drawn from the analysis.

2. Main Text

2.1. Identified Risks of Data Visualization in the Literature

One countermeasure we are taking to minimize risks is to review literature on best practices in the field for avoiding bias. Another is to build a visualization specified precisely to our user's needs. Given that our ideal end user is an AI analyst in academia with a relatively high technical competency, the cognitive risks are of highest interest to us. The social and emotional risks would be a higher priority in use cases where the visualization was seen by the general public or people from diverse backgrounds.

2.1.1 Perception Bias

Many research efforts are focused specifically on perception bias which studies the various biases and assumptions which viewers may unintentionally experience when analysing visualizations and which could obfuscate or even counteract the intentions of the visualization [8][5][20][13][21][17][11]. Outside of this research, the use of data is frequently presented as a tool for pointing facts or as an indicator of truth without there always being an acknowledgment of the existence of persuasion or bias at work in the visual representations. However, there is not an objective way to visualize data because data presentation induces different inferential models despite the main content [8].

Bresciani and Eppler outline data visualization risks into the causes(designer or user), the effects(cognitive, emotional, and social), and potential countermeasures (remedies). For example, a designer induced cognitive risk could be low accuracy due to the visualization depicting information less precisely than numbers and tables. A user induced cognitive risk could be the potential change blindness where important changes in graphics and pictures may go unnoticed by the viewer.[6]. They also suggest considering the possible negative impacts to a user: confusion, distraction, misinterpretation, manipulation, limiting reflection, or delay [7]. Bresciani and Eppler conducted a focus group study and series of interviews that articulated common sources of error in data visualization: users may believe messages conveyed in a visualization to be more sound than they are, multiple possible interpretations, and unreasonably high user competence required [7].

Sibrel et al and and Schloss et al focus on aspects of how color and spacial mapping operate as the bridging factors between "perceptual and conceptual properties in information visualizations" [21]. With the aim of understanding how people infer meaning from visual features, studying this bias makes it possible to design information visualizations that are more effective and efficient for visual communication. Both studies note the well documented "dark is more" bias which is the tendency for viewers to infer larger quantities of the values mapped with darker colors than is accurate. Sibrel et al analyzed whether the dark is more bias was still impacting on heavily spacial data where "hot spots" are present. While the "hot spot bias" was measurable when the hot spots where the locus of larger quantities, the dark is more bias was found to be more robust including when the two where in conflict [21].

Schloss et al compared the dark bias with an "opaque is more" bias to determine how the background color of the frame around the image or of the paper itself may interact with the dark is more bias and with the viewers perception [20]. They found that while both biases to exists independently, they work together in lighter background and counter in darker backgrounds which suggests "it is beneficial to use colormaps that will not appear to vary in opacity on any background color and to encode larger quantities in darker colors."

Golebiowska Coltekin analyzed if the prevalence of rainbow color schemes was based in any interpolating advantages, and found that while there are contexts which rainbow coloring may facilitate mapping, it is not intuitive and harms performance of tasks that require ordering the colors. It should therefore not be a default color mapping scheme and should only be used intentionally in specific contexts. However even in those contexts in which rainbow coloring is not a detriment to cognitive inference, it can be inappropriate for people with color vision deficiencies [11].

Independent of any any specific color scheme, Lin Thornton demonstrated that beauty, independent of data quality and graph misleadingness, increases trust in graphs across sources: scientific papers, news, and social media [17]. Though not specifically focused on beauty, Hehman Xie detailed recommend steps and procedures for individuals whose goal is to communicate patterns in data as clearly as possible to other consumers of science [13].

2.1.2 Selection Bias

Visual analysis is prone to a variety of selection bias effects, especially for high-dimensional data where only a subset of dimensions is visualized at any given time [3]. Both Borland et al and Gotz et al recognize the potential for selection bias and the dearth of accompanying visual tools to mitigate it.

Gotz et al focuses on the selection bias resulting from a critical mismatch between the very large number of dimensions in many complex real-world datasets and the much smaller number of dimensions that can be concurrently visualized. This gap in dimensionality can place a user at risk of hidden selection bias during exploratory data selection tasks. They propose an Adaptive Contextualization (AC) to mitigate this effect. "The AC approach captures a model of users' visual data selection activity, computes metrics over that model to quantify the amount of selection bias after each step, visualizes the metric results, and provides interactive tools that help users detect and assess the sources of bias as they emerge." [12]. They concluded that their idea held merit and was effective but needed further study in more multi-dimensional settings.

Bordland et al identify the risk of selection bias as even higher when analysts dynamically apply filters or perform grouping operations during ad hoc analyses [4] and these type of analytical features are a major attraction to interactive data visualization tools. Countering the effects of selection bias via bias mitigation is typically left for the user to accomplish as a separate process. Bordland et al insteas propose "Dynamic Reweighting" as a means for mitigation that helps users craft bias-corrected visualizations by "including a series of visualization designs and statistical re-weighting methods that together enable the creation of bias-corrected visualizations." Their results found that these tools were both useful and desired in workflow by their test users but it would benefit from additional simplification to make it more accessible.

Wall et al observe that most existing strategies for minimizing or mitigating cognitive bias rely on non-technical approaches ie training courses. In their work, they categorize visual analytic tools which can be used to mitigate bias and propose a design space which can be used as a template for future visualization systems to proactively guide and integrate the cognitive and analytic processes of the user [23].

2.2. Applications of Data Visualization in the Literature

In a variety of business, research, and media sectors, data visualization was applied both as a communicative tool, a way to reduce bias from various sources, and a way to increase industry productivity [2]. Banasiewicz reviewed the increased use of charts and graphs in tax auditors workflow and a tendency to over validate and over emphasise the data presented in this form. They analyzed how the figures where being used and proposed a series of accompanying visuals to help mitigate bias as well as a data visualization bias training for this specific industry [1].

They explore how current the production-focused practices in data visualization have upstaged efforts to learn about how these visualizations are consumed and actually used in organizational decision-making [1].

In the context of this project, we attempt to predict the applications of our data visualization by generating user stories that contain use cases. When we design our visualization around these user stories, the visualization can hopefully be used to fill the current gap: supporting domain specialists to learn about trends in irresponsible AI instances.

2.3. Research Approach and Best Practices

Data visualization research is research in how to best communicate about a dataset given the user of the visualization. It balances best practices in meeting user needs, with innovation that can drive new data insights. Figure 1 outlines our research flow in data visualization. We begin with the dataset and a description of our user. Knowledge of the dataset helps us know what variables will be available to generate visuals with the data. Knowledge of our ideal user helps us understand what features and how much display flexibility they will need, as well as how much domain domain knowledge and technical comfort they will have. For example, a tool built for a layperson would highlight different features than a tool built for a domain expert.

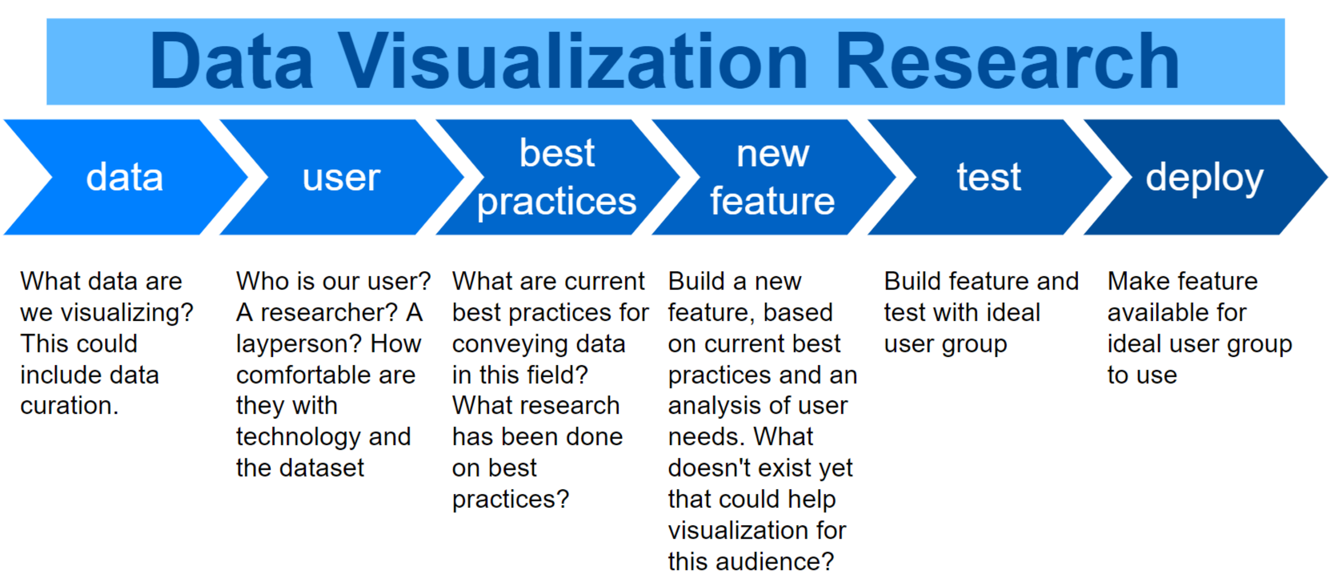

Figure 1. Research in data visualization

Our dataset consists of datapoints collected in the 'awful-ai' database, as well as the 'incidentdatabase.ai'. Each tuple contains a news article link and description of the incident, the date of occurrence and date the public found out, the affected population, actual impact and potential impact, the application of the AI, the area (industry), class and subclass labels (see figure 2), physical and company locations, additional tags, and information on researcher and news outlet.

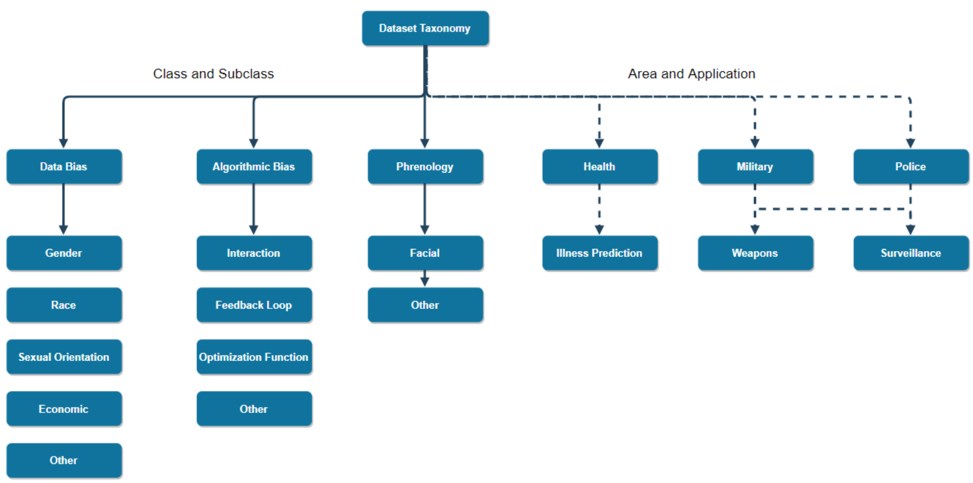

Figure 2. Taxonomy of classes, subclasses, areas and applications in our current dataset.

Our envisioned end user would be a domain expert in the field of AI ethics. We can identify what features we might need by creating user stories about our user and articulating specific ends she may be trying to accomplish through data analysis. Here is an example of some of the user stories we explored:

Research Questions: Can an Interactive Visualization Tool better support the analysis of data from instances of irresponsible AI?

Some usage scenarios:

- Rachel, a phd student who does research to support AI accountability organizations, wants to be able to view the descriptions of an incident and company responsible, associated with a scatterplot point by clicking on it.

- Rachel wants to look at the instances of irresponsible AI on a timeline, in order to determine if the impact on specific marginalized populations are growing with time.

- Rachel wants to be able to select a small section in the timeline, and zoom in to look at instances of irresponsible AI just in that timeframe, for example, looking at specific dates associated with presidential terms.

- Rachel is an AI researcher who wants to compare the domestic versus international impact of US companies irresponsibly using AI (we are trying to come up with any useful use cases that are map specific, or at least using location data.

- Rachel wants to use our tool to see if there is any trend in the following question: do global north companies misusing AI disproportionately impact global south areas?

As we explore these features to support these use cases, we consider tradeoffs with other features that are considered. A good example for weighing tradeoffs is a map feature. When we display information on a map, we are using our X and Y axis to display geographical information. These X and Y axis are one of our most valuable expressive tools in data visualization, so before displaying a map, one should consider whether the message to be communicated about the data actually requires a map. If a map is not necessary, it would be a better to reserve these X and Y axis to highlight a more precise point about the data.

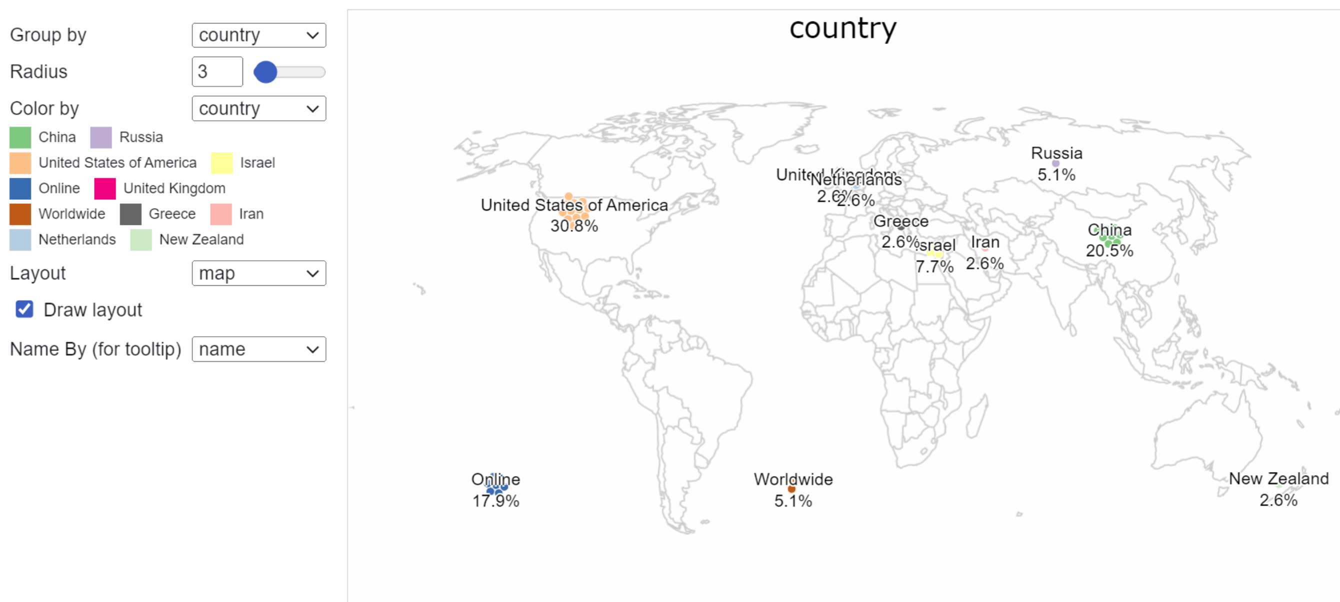

Figure 3. Current AI atlas visualization

2.4. Proposed Features

Taking our end user into account, we would want to build the following features into our data visualization. These are also organized in the table after this list, paired with the information resource that inspired this feature.

- Customizability: because our hypothetical user is knowledgeable about the domain, we trust that their assessment of what features are most important is actually more accurate than ours. Thus, incorporating high customizability would be a priority feature for this visualization. This would mean they can select which data to display and how to display it. If our user was a layperson with limited domain knowledge, we would likely give more restrictions to the customizability.

- Display AI Creators: One feature that we deemed as important from our own research about the AI field is to make sure our data visualization is capable of clearly displaying the parties responsible for each instance of AI. Our rationale behind this feature is that AI is not capable of ethical reasoning on its own and ultimately is created by humans. One way to foster more ethical use of AI may be to ensure that who created which AI instance stays part of the conversation.

- Timeline: We want the ability to sort events on a timeline, and to mark multiple important time stamps for each event. This feature was inspired by the fact that our data tracks these times, and our research user requirement example: ”our user wants to look at the instances of irresponsible AI on a timeline, in order to determine if the impact on specific marginalized populations are growing with time.”

- Sort by Class and Subclass: This feature was inspired by the classification taxonomy available in our data (See Figure 2), as well as to support user requirements similar to: ””our user wants to look at the instances of irresponsible AI on a timeline, in order to determine if the impact on specific marginalized populations are growing with time.”

Feature	**Rationale**	**Resource**
Customizability	We are not AI analysts, so ultimately we should give the analysts as much control as possible over what is displayed	Cascade.AI conference
Display AI creators	AI is not capable of ethical reasoning and the accountability must go to whoever created it	Cascade.AI conference
Timeline	A key feature of our data is the time stamps of when the incidents happened and when the public found out	Our data
Ability to sort by class and subclass	Our data is organized into classes and subclasses, which tell us what field the irresponsible AI incident was in	Our data

While there could be other additional features, and tradeoffs to weigh, these proposed features are a starting point for how we would design our data visualization.

3. Conclusion

Much of our work in this research project was skill-building in order to better understand the tools of the craft, and doing general research about the ethical AI conversation and general topics in data visualization. While we worked very closely for most of this research project, our individual contributions are outlined below.

Nolen spearheaded the bulk of the research for best practices, the literature review, curated the subclass and class of incident for the dataset, and some of the data-visualization specific context that our paper and work is situated in.

Sommer created most of the images for our technical paper, wrote the research approach, motivation and identified risks section, and curated the date of incident for much of the dataset.

4. Acknowledgements

We would like to express our sincerest thanks to our research advisor, John Alexis Guerra Gómez, our program mentor, Tom Williams, and the DREAM program team at Colorado School of Mines.

References

[1] A. D. Banasiewic. Evidence-based decision-making: How to leverage available data and avoid cognitive biases. *Routledge, Taylor Francis Group.*, 2019.

[2] E. Bendoly. Fit, bias, and enacted sensemaking in data visualization: Frameworks for continuous development in operations and supply chain management analytics. *Journal of Business Logistics*, 37:6–17, 3 2016.

[3] D. Borland, W. Wang, J. Zhang, J. Shrestha, and D. Gotz. Selection bias tracking and detailed subset comparison for high-dimensional data. 6 2019.

[4] D. Borland, J. Zhang, S. Kaul, and D. Gotz. Selection-bias-corrected visualization via dynamic reweighting. 7 2020.

[5] J. V. Braga and T. B. P. Silva. Revista brasileira de design da informação / brazilian journal of information design são paulo — v.

[6] S. Bresciani and M. J. Eppler. The risks of visualization a classification of disadvantages associated with graphic representations of information, 2008.

[7] S. Bresciani and M. J. Eppler. The risks of visualization. *Identität und Vielfalt der Kommunikations-wissenschaft (2009)*, pages 165–178, 2009.

[8] S. Campbell. The rhetoric of pathos in visualization, 2018.

[9] M. F. Dahlstrom. Using narratives and storytelling to communicate science with nonexpert audiences. *Proceedings of the National Academy of Sciences*, $11(\text{supplement}_4) : 13614 - -13620, 2009$.

[10] P. Fox and J. Hendler. Changing the equation on scientific data visualization, 2 2011.

[11] I. M. Golebiowska and A. Coltekin. Rainbow dash: Intuitiveness, interpretability and memorability of the rainbow color scheme in visualization. *IEEE Transactions on Visualization and Computer Graphics*, 28:2722–2733, 7 2022.

[12] D. Gotz, S. Sun, and N. Cao. Adaptive contextualization: Combating bias during high-dimensional visualization and data selection. volume 07-10-March-2016, pages 85–95. Association for Computing Machinery, 3 2016.

[13] E. Hehman and S. Y. Xie. Doing better data visualization. *Advances in Methods and Practices in Psychological Science*, 4(4):1–118, 2009.

[14] A. Hundt, W. Agnew, V. Zeng, S. Kacianka, and M. Gombolay. Robots enact malignant stereotypes. pages 743–756. Association for Computing Machinery (ACM), 6 2022.

[15] L. F. Klein. Feminist data visualization, 2020.

[16] S. Laato, M. Tiainen, A. N. Islam, and M. Mäntymäki. How to explain ai systems to end users: a systematic literature review and research agenda. *Internet Research*, 32(7):1–31, 2022.

[17] C. Lin and M. Thornton. Fooled by beautiful data: Visualization aesthetics bias trust in science, news, and social media, 2021.

[18] A. V. Pandey, A. Manivannan, O. Nov, M. Satterthwaite, and E. Bertini. The persuasive power of data visualization, 2014.

[19] I. D. Raji, A. Smart, R. N. White, M. Mitchell, T. Gebru, B. Hutchinson, J. Smith-Loud, D. Theron, and P. Barnes. Closing the ai accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pages 33–44, 2020.

[20] K. B. Schloss, C. C. Gramazio, A. T. Silverman, M. L. Parker, and A. S. Wang. Mapping color to meaning in colormap data visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25:810–819, 1 2019.

[21] S. C. Sibrel, R. Rathore, L. Lessard, and K. B. Schloss. The relation between color and spatial structure for interpreting colormap data visualizations. *Journal of Vision*, 20:1–20, 11 2020.

[22] U.N. Making data meaningful part 4: How to improve statistical literacy: A guide for statistical organizations. *In United Nations (Vol. 5)*, 2014.

[23] E. Wall, G. Tech, J. Stasko, and A. Endert. Toward a design space for mitigating cognitive bias in vis, 2019.