Translating Scientific Graphics for Public Audiences

Xiaoli Qiao*
University of Washington

Jessica Hullman*
Northwestern University

Abstract

It is universally acknowledged that the public need to understand and trust science to make informed decisions in everyday life. Therefore, effectively communicating scientific work to the general public is of great importance. Though scientists have developed techniques for communicating among each other through scientific writing and diagrams, it remains rare for many scientists to have to communicate to the public directly. To help scientists and others design understandable graphics for the public, we analyze 20 diagrams that represent journalists' redesigns of scientific graphics targeted at a general audience. We present 20 design strategies across five categories of techniques — Simplification, Optimizing the Visual Presentation, Adding Explanation, Making Information Relatable, and Miscellaneous — that can be applied in creating or translating scientific graphics for public consumption. By identifying specific techniques, our work demystifies the design process of understandable graphics and suggests the possibility of effectively creating such graphics with minimal expertise.

Keywords: Scientific visualization, visual analytics, visualization design and evaluation methods.

Index Terms: H.5.m [Information Interfaces and Presentation (e.g., HCI)]: Miscellaneous

1 Introduction

As the general population's trust in science declines (Edelman 2017), effective communication of science deserves more attention than ever. Science in the news often serves as the general public's primary source of scientific knowledge. In such news, journalists summarize scientists' work and sometimes borrow diagrams from publications to visually explain the topic. In many cases, journalists must modify the original diagrams, which, intended to communicate among scientists, require too much expertise for the public to understand. To scale such "graphical translation" possible for authors besides journalists, a deeper understanding of the mechanisms that are used by the journalists to translate scientific graphics for public consumption would be beneficial. Scientists could gain more systematic guidance on how to construct accessible diagrams to communicate their work, for example, or even new design tools that embody effective graphical translation techniques.

We take a closer look at original scientific visualizations and diagrams and the media versions of the same graphics and compare the differences to derive design principles. We conduct a qualitative analysis of 20 examples of scientific visualizations and diagrams that have been made more accessible by journalists. Our analysis summarizes different categories of translation techniques including adding explanation, simplification of language or interpretation, optimization of visual presentation, and making information more relatable. We describe how these design strategies at times conflict with conventional visualization design guidance like the data-ink maximization ratio or need to avoid embellishments (Tufte 1983). Our work contributes to helping authors design more understandable diagrams. We reflect on how our results could be carried out by future work to systematically render understandable scientific graphics and coach scientists to communicate their work in an accessible manner.

* xiaoliq@uw.edu
* jhullman@northwestern.edu

2 Related Work

Our study is inspired by Jonathan Corum's talk about designing science graphics for audiences with different interests and backgrounds. In his talk, Corum shares many examples where he translated graphics in scientific papers for the general public. He suggests several design guidelines, such as to tell the audience a visual story, focus the reader's attention on key aspects of the science, and use plain language wherever possible (Corum 2018). Corum's talk also addresses the research process by which he completes a graphical translation, which inspired us to undertake a study of design techniques and processes commonly employed by other journalists in translating graphics.

In understanding the simplification of scientific diagrams, Hullman and Bach (2018) analyzed graphical abstracts (GAs), graphical versions of scientific text abstracts designed to be more accessible to non-experts, to understand what visual and textual techniques are used to express scientific research. They collected and coded 54 GA samples, and identified the design patterns of four aspects - layout, depiction of time, text usage, and representational genre. Aiming to identify how scientific graphics are translated for a general audience, our work examines examples from newspapers and the Internet, which are based on but also distinctly different from the original graphics in scientific publications. Current works about diagrammatic communication usually focus on "professionally created diagrams" (Hullman, 2018) such as textbook diagrams and maps. These visualizations are created by experts in the domain, assuming the audience will have corresponding technical knowledge to understand them. Compared with prior studies, our work presents a unique perspective in that our focus is on the differences between scientific diagrams and revisions aimed at a broader audience, as well as the translation techniques that they suggest, by scientific diagrams that are intended for a broader audience.

3 Design Guidelines for Translating Graphics

We identified a set of guidelines for characterizing design techniques for translating scientific graphics.

3.1 Method

3.1.1 Sample

Our first step was to collect representative samples of journalists' efforts to re-design original scientific graphics to cater to the general public. The majority of the samples we gathered come from two sources. First, we examined into scientific journalists' work and identified explanatory illustrations associated with specific

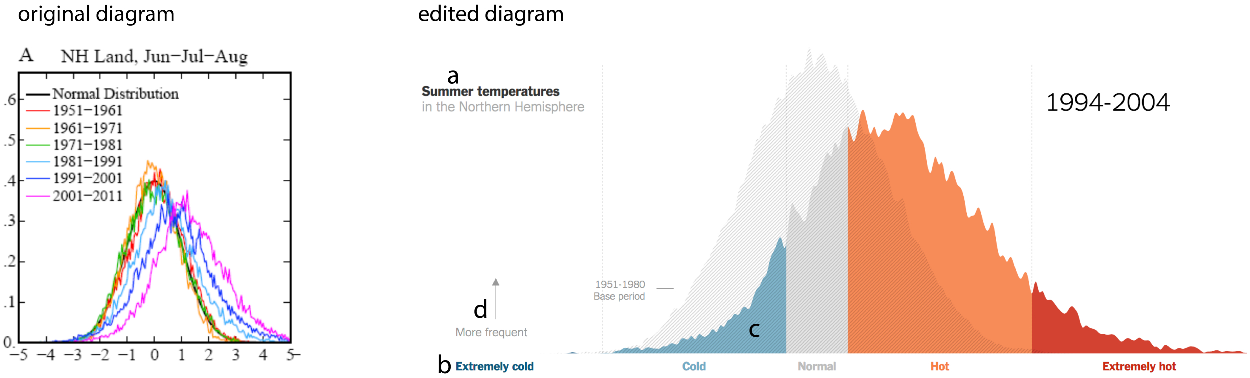

Figure 1: Original scientific graphic on northern hemisphere summer temperature (Hansen et al. 2012; left) and an edited version from the New York Times (Popovich and Pearce 2017; right).

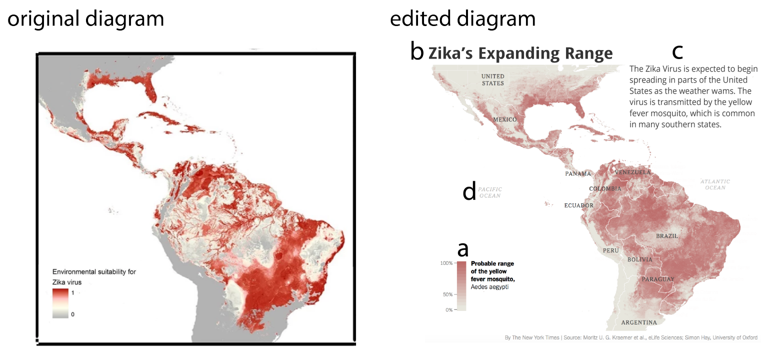

Figure 2: Original scientific graphic on environmental suitability for Zika virus (Messina et al. 2016; left) and an edited version from the New York Times (Mcneil et al. 2017; right).

Figure 3: Segment of a scientific illustration from the New York Times (Chang 2002)

scientific papers (which were typically referenced, at least through a primary author name, as part of the graphic's provenance information). After collecting a handful of samples from scientific websites such as Scientific American and Science, we browsed through The New York Times' Science section online archives to find more examples. To ensure our samples are recreations of original scientific graphics, we limited our samples to graphics that were evidently derived from the original graphics. We ended up with 20 pairs of graphics associated with various fields, ranging from public health to social science.

Next, we placed the two versions side by side to visually compare the changes. We adopted an open-ended coding approach in which each author independently analyzed the same subset of samples. After analyzing the first subset, we labeled the design strategies and created a basic taxonomy. We then applied the taxonomy to the rest of our samples, adding new labels when necessary, and resolving disagreements. After every sample pair had been analyzed, we examined our taxonomy, removed redundant labels and grouped strategies with similar design purpose. We classified the strategies found from our samples into four initial categories: **Simplification**, **Optimizing the Visual Presentation**, **Adding Explanation**, **Making Information Relatable**. We eventually included an additional category of **Miscellaneous** to account for design strategies that did not fit into any of the four basic categories.

3.2 Strategip1es with Examples

3.2.1 Simplification

Simplification can take two forms. One is to **simplify the language**. For example, journalists often **spell out acronyms** because while acronyms may be familiar to people in the specific field or expected in scientific publications, their meanings tend to be obscure to a general audience (e.g., in Figure 1a "NH" is expanded into "Northern Hemisphere" to assist understanding.). Other strategies of simplifying language include to **rewriting existing labels in simpler language** (e.g., "periarterial venous rete" was rephrased to "net of veins around artery") and **adding a description of the unit or scale** (e.g., in Figure 1b, the numeric x-axis is translated into descriptive words such as "extremely cold" and "extremely hot").

The other aspect is to **simplify the interpretation**. To explain a complex concept or to show trend over time, journalists usually **break the visualization into steps**. For instance, Figure 1, originally an animated visualization on the website, shows the change in temperature by decade to provide a more concrete impression of the changes. **Changing the measurement** is another common technique to simplify interpretation. In one example, "Environmental suitability for Zika virus" is changed into "Probable range of the yellow fever mosquito" because the latter is easier to understand (Figure 2a). To increase data-ink ratio, journalists often **intentionally leave out some information** that is not associated with the target pattern or concept and **remove visualization of uncertainty** (such as by removing error bars or descriptions of margins of error). Some of these techniques are described in prior work on narrative visualization (Segel and Heer 2010, Hullman and Diakopoulos 2011), but simplification is often not addressed directly in more canonical examples of visualization guidance (e.g., Tufte 1983, Cleveland 1985).

3.2.2 Optimizing Visual Presentation

We found that journalists often make direct changes to the graphics to improve the visualization. For instance, journalists sometimes **change the colors** of the original graphics for a more attractive and distinguishable look, such as using more pleasing palettes. They also tend to **change the visual encoding** when the scientific graphic has unorganized colors and fails to convey the information effectively (e.g., in one of our samples pairs, the researchers used a diverging color palette to show temperature from low to high, but the journalist, in the edited version, changed it to a simpler sequential color scheme). Other strategies employed by journalists to optimize the visual presentation are to **add a visual encoding** (e.g., blue, gray, orange and red are added as a redundant encoding in Figure 1c to denote 'cold', 'normal', 'hot', and 'extremely hot'), **add a visualization of data** (e.g., visualize data originally in a table) and **change orientation or ordering** of parts of a figure, typically for aesthetic reasons.

3.2.3 Adding Explanation

We observed two examples of strategies for adding explanation - **adding explanation of the content** or **adding explanation of the graph** itself. When journalists want to explain the content more clearly, they **add title** (Figure 2b) and **provide background information or summary** (Figure 2c) to let the audience quickly get familiar with the topic. When journalists think it's necessary to explain the diagrams, they **add labels to part of the diagram** (which were unlabeled in the original) (Figure 2d) or **indicate how to read the graph** (Figure 1d) or **directly interpret the pattern shown in the original graph**, both typically through annotation.

3.2.4 Making Information More Relatable

We also noticed that journalists sometimes included a variety of visual embellishments in their work to make the graphic design more relatable to the audience. One approach is to use **more realistic illustrations**. For instance, viruses and bacteria in the scientific papers are usually drawn as circles, however, in the edited version, most journalists chose to change their appearance to be more realistic. Another strategy to make information relatable is to add **a reference to a familiar object**, such as adding an icon of a person, building, or other object as a reminder of the domain, or as a size reference (Figure 3). Adding pictorial objects purely to support recognition of some object or entity conflicts with canonical guidance like the data-ink maximization ratio (Tufte 1983). This strategy is however supported by more recent studies of the impact of chart-junk or other recognizable objects in visualization interpretation (Borgo et al. 2012, Borkin et al. 2013). Further, while some work has been done in visualization to understand (Chevalier et al. 2014) and support generation of measurement analogies (Kim et al. 2016 CHI, Hullman et al. 2018), authors are generally left on their own to devise such references. The last strategy we identified as serving to make the information relatable is journalists would **provide a typical case for the audience to better understand the special case** talked about in the article. For instance, while the "Base Period" in Figure 1 was already present in the scientific version, in some other visualizations we observed, the journalist added a baseline comparison of some type. This creates a vivid contract to help the audience perceive the anomaly that might otherwise not be perceivable given their lack of expertise.

3.2.5 Miscellaneous

We observed several other design techniques that did not fit into any of the categories above. Sometimes journalists **added or changed information groupings from the original diagram**. As shown in Figure 1, editors slightly changed the grouping of years. Another frequently used approach is to **add new data or information**. For example, in one diagram about labor force participation rates, the journalist added global data to show a holistic view while the original source only has domestic data.

4 Discussion and Future Work

Our analysis suggests that journalists and others who translate scientific visualizations and diagrams for broader audiences have identified specific techniques that serve graphical translation work. These strategies appear to be influenced by principles that run orthogonal to most canonical visualization design guidance, which focuses on the clarity of the encodings but does not necessarily address how to scaffold understanding through textual and graphical elements. While we observed some techniques, such as *changing the measurement*, only a couple times, other techniques, like *providing background information or summary*, appeared in 63% of visualizations in our sample. This implies that journalists use a common set of strategies.

One avenue for future work is to try to teach scientists these principles. Paired examples like those in our sample could be used to demonstrate how aspects of an original scientific graphic can be systematically revised to increase the clarity of language, interpretation, visualization, etc. It is possible that short workshops could be used teach the techniques to scientists through redesign activities, similar to how one might teach lay people to become better critics of information graphics.

Another goal for future work is to embed techniques like those that we observed into design assistants and systems. For example, given the full text of a scientific article and lexical simplification techniques (e.g., Kim et al. 2016), it may be possible to auto-generate background text or titles for a scientist attempting to redesign a graphic for the public. Crowdsourcing may also be a promising way to solicit feedback on a scientist's attempt at a more accessible diagram, such as if crowdworkers comments can be classified according to the taxonomy we present to provide focused feedback on what to improve. Through algorithms that employ OCR, computer vision, and lexical simplification, it may be possible to build human-in-the-loop design systems made expressly for scientists communicating their work.

References

[1] Cleveland, W. S., & Cleveland, W. S. (1985). *The elements of graphing data* (Vol. 2). Monterey, CA: Wadsworth Advanced Books and Software.

[2] Corum, J. (2018). Design for an Audience (presented at University of Copenhagen Faculty of Health Sciences). http://style.org/ku/

[3] Edelman (2017). Edelman Trust Barometer. http://edelman.com/trust2017

[4] Hansen, J., Sato, M., & Ruedy, R. (2012). Public perception of climate change and the new climate dice. arXiv preprint arXiv:1204.1286.

[5] Segel, E., & Heer, J. (2010). Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics*, *16*(6), 1139-1148.

[6] Hullman, J., & Bach, B. (2018, June). Picturing Science: Design Patterns in Graphical Abstracts. In *International Conference on Theory and Application of Diagrams* (pp. 183-200). Springer, Cham.Hullman, J., & Diakopoulos, N. (2011). Visualization rhetoric: Framing effects in narrative visualization. *IEEE transactions on visualization and computer graphics*, *17*(12), 2231-2240.

[7] Kim, Y-S., Hullman, J., Burgess, M., and Adar, E. (2016). SimpleScience: Lexical Simplification of Scientific Terminology. *Proc. EMNLP 2016*.

[8] Krueger, A. B. (2017). Where Have All the Workers Gone?: An Inquiry into the Decline of the US Labor Force Participation Rate. *Brookings papers on economic activity*, *2017*(2), 1-87.

[9] Popovich, N., and Pearce, A. (2017). It's Not Your Imagination. Summers Are Getting Hotter. The New York Times. https://www.nytimes.com/interactive/2017/07/28/climate/more-frequent-extreme-summer-heat.html

[10] Tufte, E., & Graves-Morris, P. (2014). The visual display of quantitative information; 1983.

[11] Borgo, R., Abdul-Rahman, A., Mohamed, F., Grant, P.W., Reppa, I., Floridi, L. and Chen, M. (2012). An empirical study on using visual embellishments in visualization. *IEEE Transactions on Visualization & Computer Graphics*, (12), pp.2759-2768.

[12] Borkin, M.A., Vo, A.A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A. and Pfister, H. (2013). What makes a visualization memorable?. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), pp.2306-2315.

[13] Kim, Y.S., Hullman, J. and Agrawala, M. (2016, May). Generating personalized spatial analogies for distances and areas. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 38-48). ACM.

[14] Hullman, J., Kim, Y.S., Nguyen, F., Speers, L. and Agrawala, M. (2018, April). Improving Comprehension of Measurements Using Concrete Re-Expression Strategies. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 34). ACM.