Did you see that? The role of graphs in journalism

Chantelle Ivanski & Stephanie Raposo

Final Research Paper

Data Visualization, York University

Dr. Michael Friendly

June 28, 2021

Did you see that? The role of graphs in journalism

Data journalism has been on the rise worldwide since the 1990s (Young et al., 2018), yet its practicality for readers still seems to be quite mixed. Providing clear and effective data visualizations in journalism, either to supplement or replace written content, can serve many useful functions. For example, visualizations can be used to tell a story or answer questions (when proper context is given), personalize the display and interactivity of data for readers, and facilitate stronger and more transparent connections between authors and readers (e.g., McBride, 2016; Rodríguez et al., 2015). In fact, during the COVID-19 pandemic, data visualizations in the media in Korea were reported to be twice as high as articles that were unrelated to the pandemic (Kwon et al., 2021), thus suggesting how useful visualizations can be for illustrating important information. However, despite these convincing numbers, Kwon et al. (2021) also found an evident discrepancy between the level of literacy required for interpreting graphs in typical school curriculums, compared to in the media. Moreover, graphs in media can confuse readers if the wrong type of graph is displayed (Glazer, 2011) or if visuals lack context, which can lead to biased interpretations of data from both the authors and readers (an issue that is not unique to pandemic-related journalism; Kwon et al., 2021). Given that data visualizations in journalism are on the rise and are associated with mixed outcomes for readers, how might we better understand their practicality and strengthen visualizations to make them more easily (and properly) interpreted? The proposed research, therefore, aims to explore the role of colour and interactivity as two key features that can impact understanding. Although research has explored the extent to which data visuals boost or detract from journalism overall, it is not clear how much readers *pay attention* to data visualizations in their everyday lives, or whether integrating (or combining) certain features makes a difference. To address these questions, we will explore the experiences

of people viewing data visualizations that range in how colourful and interactive they are, which will also allow us to identify whether these features play a role in how people focus on, interpret, and understand key messages in journalism.

The Function and Presentation of Data Visualization in Journalism

The increasing use of data visualizations in journalism may be, in part, driven by how functional they are for helping authors form stories. According to Rodríguez et al. (2015), data visualization can provide insight to readers by helping them to understand the "big picture" message, giving them options to restructure the data, presenting patterns among the data, and enabling them to create mental images of the visuals. In fact, visuals can even be more effective than written text because they tend to provide clearer and more accessible explanations of the data (Weber et al., 2018). This interplay between visuals and written content is otherwise known as the "text-visualization relationship," which describes the intentional placement of visuals within an article to either lead a story (e.g., if placed at the beginning) or supplement written text (e.g., alternating visuals and text; Weber et al., 2018). Overall, data visualizations can play a key role in shaping how readers digest information, but this might also depend on the structure in which authors choose to tell their story.

Data visualizations can be presented in the form of 1) the "martini glass structure," an author-driven method in which the author controls the narrative and allows readers to interact with the data at the end; 2) an "interactive slideshow" which is both author- and reader-driven wherein the author presents data as a slideshow and allows readers to interact with data before moving to the next stage when they are ready; or 3) a "drill down story" which is reader-driven in that the author predetermines the data and interactive elements, and the readers have more autonomy to interact with the data at any time (Segal & Heer, 2010). As such, authors can decide

on the most effective ways to present their information with visuals, but once the articles are out of their hands, how practical are they for readers?

The Practicality of Data Visualization in Journalism for Readers

Given the prevalence and function of data visuals in journalism today (Monteiro & Ainley, 2004), it is imperative that we also consider how *useful* graphs in journalism are from the reader's perspective. A common misconception is that key messages of visualizations in journalism are clear and obvious to readers (Ainley, 2000; Monteiro & Ainley, 2004); however, this is not always the case, and there is a lack of consistency between graph interpretation in educational settings versus everyday life (Kwon et al., 2021; Watson, 1997). In fact, the ways in which we can apply our literacy skills (i.e., understanding the key takeaways of visuals, making comparisons within the data, and extrapolating beyond the data) tend to be limited to the most basic graphs that we are formally taught to interpret (Curcio, 1987) and may not transfer to more complex or unusual graphs that are also portrayed in journalism (Monteiro & Ainley, 2004; Watson, 1997). To interpret graphs more effectively, readers can strive to continue developing a "critical sense," which is the ability to simultaneously balance our criticism of graphs with our own beliefs, emotions, and knowledge (Monteiro & Ainley, 2004).

Another way that the practicality of data visualizations depends on readers' perspectives and understanding of graphs is how two key features (i.e., colour and interactivity) are applied to visuals to facilitate more effective knowledge translation. Importantly, readers tend to overlook unattractive graphs and perceive stories as more reliable when visual features are carefully selected (see Weber et al., 2018). The remainder of this paper will focus more closely on visual appeal in journalism from the readers' perspective.

Colour in Data Visualization

In data visualizations, colours are important for helping to draw readers' attention and engagement (MacDonald, 1999). However, colour usage is only valuable when it is done correctly. When applied incorrectly to visuals, colour can be distracting to readers and may result in further misunderstanding and confusion (Few, 2008). For example, rainbow colours are often used in graphs, but they have been shown to be confusing to interpret (Li, 2013). This is especially true for colourblind individuals, with red-green colourblindess being one of the most common types (Few, 2008; MacDonald 1999). It is also not as straightforward to order rainbow colours as it is to order a single colour in multiple shades, especially for sequential data (Few, 2008; Li, 2013). Even if the data is not impacted by sequence, however, colour needs to be carefully applied because it can be confusing if it is added simply for aesthetic reasons without consideration for how it will be interpreted (Few, 2008). Moreover, when well thought through, colour can make graphs easier to engage with for readers. For example, a single colour palette that uses softer colours is easier on the eyes and simple to interpret (MacDonald 1999; Few, 2008). Thus, given the import role colour can play in interpretation and engagement, we are including it as one of the main variables of interest in the proposed study.

Interactivity in Data Visualization

Interactivity—the extent to which readers can control and explore the information they view in a graph (Young et al., 2018)—is another key feature of data visualizations that should be prioritized and planned carefully in journalism. Beginning as a tool for readers to navigate graphs (e.g., zooming in/out, filtering by specific information, clicking points for additional detail, reconfiguring data into readers' preferred formats), data journalism has evolved to include more interactive options, such as sharing to social media platforms and comment sections to discuss ideas with other readers (Aitamurto et al., 2011; Young et al., 2018). Like colour, authors

must carefully consider potential drawbacks to interactivity. For example, it can be easy to distract readers with simple yet unnecessary interactive elements, which often end up being overwhelming for readers (Burmester et al., 2010), because they are free of charge and appropriate elements would either require more advanced skills or collaboration with third parties (Young et al., 2018; Zion & Craig, 2014). Despite these considerations, interactive elements are crucial in journalism as they encourage readers to spend more time exploring and engaging with articles, they empower readers to access more personally relevant information, they improve readers' recall of key information, and they create more inclusive environments for readers with varying literacies (Aitamurto et al., 2011; Dada, 2016; Young et al., 2018). Taken together, interactivity is a critical way to facilitate better data journalism (Aitamurto et al., 2011), and as such, it is the second main variable of interest in our proposed study.

The Proposed Study

Data visualizations in journalism play a crucial role in conveying important information to readers, which has been shown to be largely driven by features like colour and interactivity. Despite how aesthetic and functional authors aim to make their visuals, it is ultimately the reader who decides whether to focus on key takeaways from the visuals, written text, or a combination of both. In our proposed research, we will conduct an experimental study to learn more about readers' experiences when presented with various formats of data visualizations in journalism. Specifically, the proposed study will assess the following two questions: 1) when people consume journalism, do they pay more attention to graphs or written content, and 2) do features of graphs (i.e., colour and interactivity) influence whether people focus more on graphs or written content, as well as how accurately they interpret the main message?

Method

Participants

A representative sample of 500 Canadian adults will be recruited through Qualtrics

Panels to participate in this study. Participants will be compensated with $2.50 CAD and will be

excluded if they are under 18, guess the purpose of the study, or skip questions about our key

dependent variable (see below).

Measures

Study Advertisement. Potential participants will see an advertisement which will

indicate that we are seeking participants for a study about birth rates in Canada. To ensure

reliability of the results, the true nature of the study will be hidden to participants until they

complete the survey.

Articles. Participants will be randomly assigned to see one of four articles about birth

rates in Canada. Each 500-word article will be accompanied by a graph embedded in the article

which will either be in (1) black and white and not interactive, (2) black and white and

interactive, (3) colour and not interactive, or (4) colour and interactive (see Figure 1).

Importantly, in all conditions, the graphs will show that the birth rate in Canada has been

increasing since the 1960s, whereas the written content will conclude that birth rates have been

decreasing.

Survey. The key dependent variable in this study is a written summary about the article

that participants will be asked to write. After reading the article participants will be asked, "In a

couple of short sentences, please summarize the main point of the article you were just shown."

Following this, they will be asked whether they paid more attention to the graph, the written

content, or a combination of both. Finally, they will also be asked a few questions about their

engagement with news articles and graphs in general (i.e., how often they read the news in
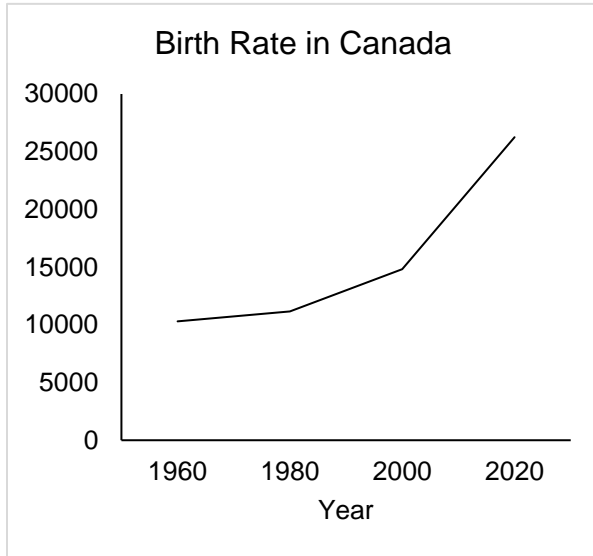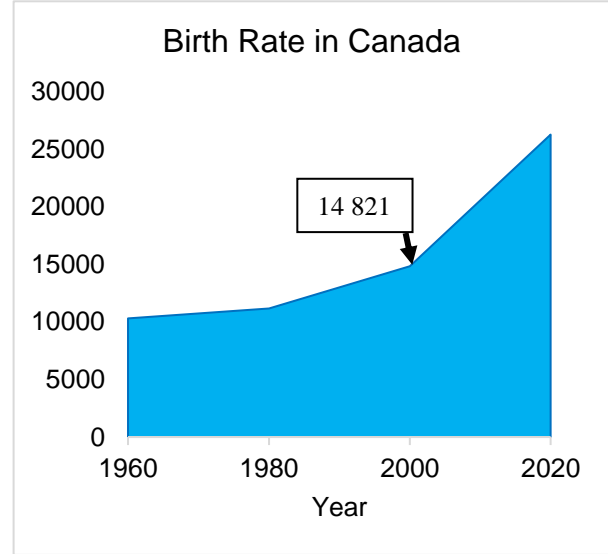
(a) Black and white, not interactive
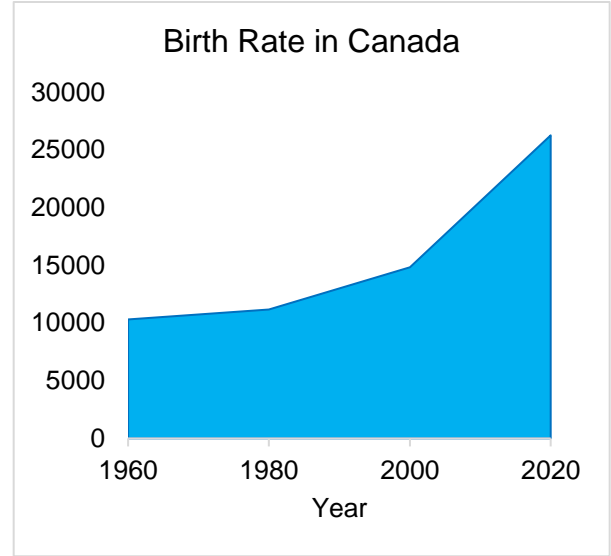
(b) Colour, not interactive
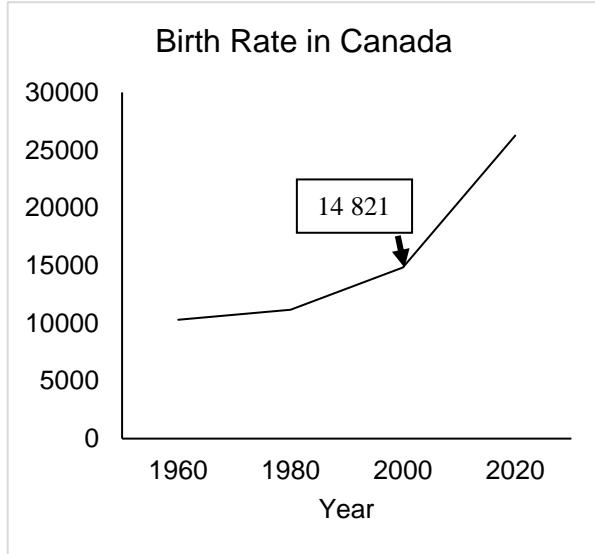
(c) Black and white, interactive

(d) Colour, interactive

Figure 1. The four graphs that will be presented with in each condition. In the interactive condition participants will be able to select different parts of the graphs and see the values appear.

general, which news sites they use most often, and how much they rely on graphs when reading

journalistic articles).

Intermixed among the questions of interest, participants will also answer questions about

birth rates and parenting in Canada, which will not be analyzed but will be used to further

disguise the purpose of the study. Specifically, they will be asked to report on what they think

the cause of change in birth rate is (e.g., available social programs) and to rate how much they

think each item is a cause on a scale from 1 = "not at all" to 5 = "entirely". They will also be

asked how the birth rate change will impact the future economy (rated on a 5-point scale from

"entirely positively" to *"entirely negatively",* with *"no impact"* in the center), and whether they

currently have, or plan to have children, and if so, how many.

Demographics. Participants will also be asked a series of demographic questions to

better understand the sample, including their age, gender, relationship status, and cultural

background.

Demand Characteristics. At the end of the study, participants will be asked what they

thought the purpose of the study was to ensure they remained blind to the true purpose. If they

guess the purpose of the study, we will remove them from analyses in case they altered their

responses to be more in line with their perceived goal of the study.

Procedure

Potential participants will be recruited through Qualtrics Panels to ensure a representative

sample. Interested individuals will then be shown the consent form and detailed information

about the study and their rights as participants. If they are still interested in participating, they

will be required to confirm their consent before continuing.

First, participants will be told that they are going to be given an article about the changing birth rates in Canada and asked to read it as they would any other journalistic article they would come across in their daily lives. They will then be given a short article describing the decrease in birth rates with one of the four graphs depicting an increase embedded within. Participants will be able to advance forward at any point after they finish with the article, or they will be auto advanced forward after three minutes. The auto advance is being implemented to ensure that people are engaging with the articles as they typically would and that they are not studying the material at length. After reading the article, they will be asked to summarize the main points of the article and then be presented with the remainder of the survey items with the camouflage questions spread throughout. They will then be presented with the demographic questions before responding to the items meant to determine if they guessed the true purpose of the study. Finally, they will be presented with a debriefing form explaining the true purpose of the study and asked to provide post-debriefing consent. Participation in this study will take approximately 20 minutes and all hypotheses, methods, and materials will be preregistered on the Open Science Framework before data collection begins.

Expected Results

Based on how participants describe the main point of the article, we will be able to identify whether they actually read through the written content or relied solely on the graph to understand how birth rates have changed. Regarding the key message of the article, we expect that participants will be more likely to report that the birth rate has been increasing in Canada as described in the graphs, rather than decreasing as described in the written portion of the article. More specifically, we anticipate that participants will be most likely to report an increased birth rate when they see the colour + interactive graph, and least likely to report the increase when

they see the not black and white + interactive graph. The graphs that are either black and white + interactive or coloured + not interactive are expected to fall in the middle. Regarding participants' attention, we expect that most people will report that they paid more attention to the graph than the written content, followed by a combination of the two, and the least number of people will report that they paid more attention to the written content than the graph.

Discussion

Implications

The use of graphs and other forms of data visualization has become more common in journalistic articles (Young et al., 2018). Although research has demonstrated what parts of graph matter most, little work has examined how much attention is paid to data visualizations compared to other elements of news articles. This work will allow us to demonstrate the importance of including graphs in journalism to help readers more easily understand the content being discussed. It will also help to show the importance of not simply including some sort of data visualizations but including *good* visualizations that are easily interpretable (e.g., using colour to enhance the visualization when appropriate). Similarly, it will demonstrate the importance of interactivity in graphs and how it can further increase the value that people get out of journalistic articles when they are able to interact with the media they are consuming. Lastly, this project will provide a valuable foundation for future research on data visualization in media.

Future Directions

The proposed work will present several promising avenues for future research to advance our understanding of how data visualizations in journalism are viewed, used, and interpreted. First, future work could consider how other features beyond (or in addition to) colour and interactivity might impact readers' attention and understanding. Second, research could extend

our proposed work to other methods, such as eye tracking procedures to measure the extent to which people actively look at, inspect, and visually compare features of graphs with written text. Third, future work could assess the generalizability of findings to more diverse audiences with varying levels of literacy and experience, and more complex visuals that are not formally taught but still presented in journalism. Fourth, while our proposed research would focus on one example of a text-visualization relationship, future research could explore the impact of other configurations of visuals and written text on readers' experiences. Finally, a critical next step would explore underlying mechanisms for why people might direct more attention to–and more successfully interpret–graphs versus written text (e.g., readers' experience/knowledge with interpreting graphs). Overall, we are optimistic that our proposed research would inspire several new lines of work that would encourage journalists to share more effective visuals and enhance readers' experiences to elevate the translation of knowledge in journalism today.

References

Ainley, J. (2000), Constructing purposeful mathematical activity in primary classrooms. In C.

Tikly and A. Wolf (eds.), *The Maths We Need Now: Demands, Deficits and Remedies*,

University of London, London, 138-153.

Aitamurto, T., Sirkkunen, E., & Lehtonen, P. (2011). Trends in data journalism. *Espoo: VTT*, 0-

27.

Burmester, M., Mast, M., Tille, R., & Weber, W. (2010). "How Users Perceive and Use

Interactive Information Graphics: An Exploratory Study." *IEEE Proceedings of the 14th

International Conference Information Visualization (IV 10)*, London, 361–368.

Curcio, F. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal

for Research in Mathematics Education, 18*(5), 382-393.

Dada, J. (2016). *Engaging to explore? An online experiment to investigate the impact of

interactivity in data visualizations.* [Master's thesis, University of Oxford].

Few, S. (2008, February). Practical Rules for Using Color in Charts. *Visual Business Intelligence

Newsletter.* Retrieved from

http://www.perceptualedge.com/articles/visual_business_intelligence/rules_for_using_colo

r.pdf

Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in

science education*, *47*(2), 183-210.

Kwon, O. N., Han, C., Lee, C., Lee, K., Kim, K., Jo, G., & Yoon, G. (2021). Graphs in the

COVID-19 news: A mathematics audit of newspapers in Korea. *Educational Studies in

Mathematics*, 1-18.

Li, A. (2013, October, 3). *Why rainbow colors aren't the best option for data visualizations.* Poynter. https://www.poynter.org/archive/2013/why-rainbow-colors-arent-always-the-best-options-for-data-visualizations/

McBride, R. E. D. (2016). The ethics of data journalism. *Digital Commons University of Nebraska-Lincoln,* 1-44.

MacDonald, L. W. (1999). Using color effectively in computer graphs. *IEEE Computer Graphics and Applications,* 19(4)*,* 20-35. doi:10.1109/38.773961.

Monteiro, C., & Ainley, J. (2004). Exploring the complexity of the interpretation of media graphs. *Research in Mathematics Education*, *6*(1), 115-128.

Rodríguez, M. T., Nunes, S., & Devezas, T. (2015, September). Telling stories with data visualization. In *Proceedings of the 2015 Workshop on Narrative & Hypertext,* 7-11.

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

Watson, J.: 1997, 'Assessing statistical literacy through the use of media surveys', in I. Gal and J. Garfield (eds.), *The Assessment Challenge in Statistics Education*, IOS and Press International Statistical Institute, Amsterdam, 107-121.

Weber, W., Engebretsen, M., & Kennedy, H. (2018). Data stories: Rethinking journalistic storytelling in the context of data journalism. *Studies in Communication Sciences*, *2018*(1), 191-206.

Young, M. L., Hermida, A., & Fulda, J. (2018). What makes for great data journalism? *Journalism Practice*, *12*(1), 115-135, DOI: 10.1080/17512786.2016.1270171

Zion, L., & Craig, D. (Eds.). (2014). *Ethics for digital journalists: Emerging best practices*. Routledge. pp. 202-221.