

Can interactivity make you think?

Increasing critical engagement with data through interactive visualizations

Tomas Petricek
University of Kent and
The Alan Turing Institute
tomas@tomasp.net

Josh Cowls
The Alan Turing Institute and
the University of Oxford
jcowls@turing.ac.uk

Open data and online news have powered the rise of data journalism, enabling reporters to publish detailed factual analyses. At the same time, public trust in the media is at historic lows. Rather than engaging with in-depth analyses, readers live in echo chambers and fall for fake news, misled by their biases¹ as public trust in media continues to decline². This suggests a curious paradox, wherein *authority* and *accessibility* tend to be negatively correlated: news that is authoritative isn't often accessible, and news that is accessible isn't typically authoritative³.

We implement and study a new way of presenting data that promotes both authority and accessibility. Using theoretical arguments and experimental evaluation, we argue that this has the potential to reverse the declining trust in the news. Our approach encourages readers to question their assumptions about data and promotes data literacy⁴. Our contributions are twofold. First, we present an open-source library of interactive data visualizations that allows novel presentation of various kinds of data. Second, we propose an empirical evaluation of how exposing readers to these data visualizations affects understanding of and trust in the information presented in news stories, as compared with text-only and conventional, static graphs.

Guess-first data visualizations

In the first part of the paper, we describe a novel library for creating interactive visualizations. Our guess-first visualizations ask the reader to make a guess before showing the actual data⁵. This encourages an active approach to data – readers need to make their assumptions about the topic explicit before they are confronted with the accurate data. To promote full transparency, each visualization is also backed by simple source code, accessible via an “open source code” link, that shows where data comes from and how it has been transformed⁶.

The library is inspired by pioneering work of newsrooms such as the New York Times⁷, but we generalise the idea to work on multiple kinds of data and charts. Three examples shown in Figure 1 include completing a time series in a line chart (by drawing the rest of the line), guessing the relative magnitudes of a bar chart (by dragging the bars or columns) and aligning a known time series to a list of known events (by dragging it up or down). The different visualizations let us expose several kinds of reader's biases including beliefs about numerical facts and causation.

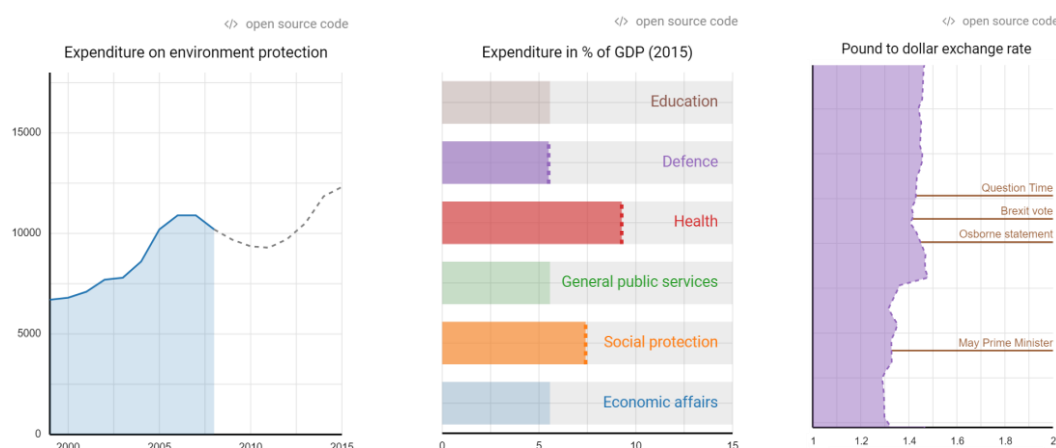


Figure 1. Three interactive “You Guess” visualizations: (left) User has to complete a line before the second part is shown; (middle) User has to guess value for each bar before correct sizes are shown and (right) user needs to align a time series with outlined events by dragging it up or down.

Experimental evaluation

In the second part of the paper, we introduce the results of an experiment to ascertain the extent to which presenting information as depicted above affects how readers understand and retain it. As part of an online experiment with participants recruited through Mechanical Turk⁸, we assign participants to one of three different stories: a text-only article; an article with a conventional, static visualisation, and the dynamic, interactive visualisation described above.

After exposing participants to the article, we use a self-reported survey to assess: (i) how well they retained the information conveyed, (ii) how engaging they felt the article to be, and (iii) how likely they would be to share the article more widely. Through these follow-up questions we ascertain whether authority and accessibility are each promoted by the “You Guess” methodology of presenting data. The experiment will therefore offer tentative evidence in relation to the use of novel forms of data presentation and interaction and its impact on how the news is understood and trusted.

References

- Bakir, V., Andrew McStay (2017) Fake News and The Economy of Emotions, *Digital Journalism*, 6:2, 154-175, DOI: 10.1080/21670811.2017.1345645
- Buchnan, L., H. Park and A. Pearce (2017). You Draw It: What Got Better or Worse During Obama’s Presidency. Last accessed 5 April 2018. Available online at: <https://www.nytimes.com/interactive/2017/01/15/us/politics/you-draw-obama-legacy.html>
- Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon. com’s Mechanical Turk. *Political Analysis*, 20(3), 351-368.
- Guess, A., B. Nyhan, and J. Reifler (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 us presidential campaign (2018). Last accessed 5 April 2018. Available online at: <https://www.dartmouth.edu/~nyhan/fake-news-2016.pdf>
- Nicolau, A. and C. Giles (2017). Public trust in media at all time low, research shows. *Financial Times*. Available online at: <https://www.ft.com/content/fa332f58-d9bf-11e6-944b-e7eb37a6aa8e> Last accessed 5 April.
- Petricek, T. (2017). The Gamma: Programming tools for open data-driven storytelling. In *Proceedings of the European Data and Computational Journalism Conference*. University College Dublin, Dublin, Ireland, 6-7 July 2017. 2017-07-06. Available at: <http://hdl.handle.net/10197/8634>
- Riederer, C., Hofman, J. M., & Goldstein, D. G. (2018). To put that in perspective: Generating analogies that make numbers easier to understand. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI’18)*. ACM, New York, NY, USA.

¹ (Guess, Nyhan & Reifler, 2018) provide an empirical evaluation and estimate that 1 in 4 Americans visited a fake news website between October 7 and November 10, 2016.

² See for example (Nicolau & Giles, 2017)

³ (Bakir & MacStay, 2017) analyse the socio-economic factors contributing to fake news and note the rise of emotionally targeted fake news – a format that is more accessible than detailed data-driven analyses.

⁴ (Riederer, Hofman & Goldstein, 2018) show how the use of data can guide better understanding. Their focus is on putting numerical values in context, but the broad approach is similar to ours.

⁵ Sample articles using the library (shown in Figure 1) can be found at <http://gamma.turing.ac.uk>. The visualizations are available as part of the open-source package TheGamma: <https://github.com/the-gamma>

⁶ We focus on interactive data visualizations here, but the transparency of data access and transformation is an equally important component of our work. This has been described in (Petricek, 2017)

⁷ See (Buchnan, Park & Pearce, 2017)

⁸ As noted by (Berinsky, Huber & Lenz, 2012), despite several disadvantages, Mechanical Turk provides a suitable method for recruiting survey participants and is widely used in current research.

Author biographies

Tomas Petricek is a Lecturer in School of Computing at University of Kent, Collaborating Fellow at the Alan Turing Institute and a recipient of the Google Digital News Initiative (DNI) Innovation Fund grant. In his PhD at University of Cambridge, he developed foundations of context-aware programming languages. His recent work has been focused on simplifying programming tools for reproducible data science and making the creation of transparent, open data analyses accessible to non-programmers such as data journalists.

Josh Cows is a Research Assistant in Data Ethics at the Alan Turing Institute and a Research Associate at the Digital Ethics Lab, Oxford Internet Institute, University of Oxford. Josh's research centers on the impact of the internet on politics and the media, and he holds graduate degrees from the Oxford Internet Institute and MIT's Comparative Media Studies program. He has studied the implications of big data, open data, state surveillance, and the use of social media in political campaigns, and he has co-authored work appearing in *New Media & Society* and *Policy and Internet*.