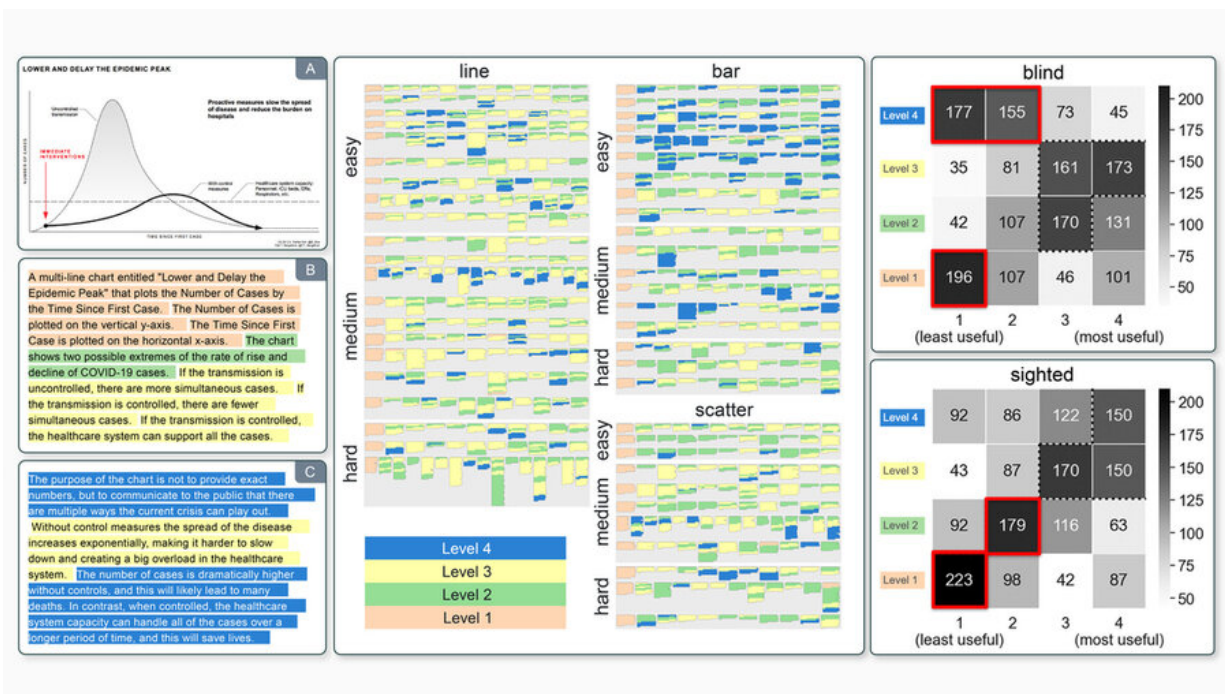# Blind and sighted readers have sharply different takes on what content is most useful to include in a chart caption

October 12 2021, by Adam Zewe

Three columns containing various graphics. The first contains the canonical Flatten the Curve coronavirus chart and two textual descriptions of that chart, color-coded according to the four levels of the semantic content model presented in the paper. The second contains a corpus visualization of 2,147 sentences describing charts, also color-coded, and faceted by chart type and difficulty. The third contains two heat maps, corresponding to blind and sighted readers' ranked preferences for the four levels of semantic content, indicating that blind and sighted readers have sharply diverging preferences. Credit: Massachusetts Institute of Technology

In the early days of the COVID-19 pandemic, the Centers for Disease Control and Prevention produced a simple chart to illustrate how measures like mask wearing and social distancing could "flatten the curve" and reduce the peak of infections.

The chart was amplified by news sites and shared on social media platforms, but it often lacked a corresponding text description to make it accessible for blind individuals who use a screen reader to navigate the web, shutting out many of the 253 million people worldwide who have visual disabilities.

This alternative text is often missing from online charts, and even when it is included, it is frequently uninformative or even incorrect, according to qualitative data gathered by scientists at MIT.

These researchers conducted a study with blind and sighted readers to determine which text is useful to include in a chart description, which text is not, and why. Ultimately, they found that captions for blind readers should focus on the overall trends and statistics in the chart, not its design elements or higher-level insights.

They also created a conceptual model that can be used to evaluate a chart description, whether the text was generated automatically by software or manually by a human author. Their work could help journalists, academics, and communicators create descriptions that are more effective for blind individuals and guide researchers as they develop better tools to automatically generate captions.

"Ninety-nine-point-nine percent of images on Twitter lack any kind of description—and that is not hyperbole, that is the actual statistic," says Alan Lundgard, a graduate student in the Computer Science and

Artificial Intelligence Laboratory (CSAIL) and lead author of the paper. "Having people manually author those descriptions seems to be difficult for a variety of reasons. Perhaps semiautonomous tools could help with that. But it is crucial to do this preliminary participatory design work to figure out what is the target for these tools, so we are not generating content that is either not useful to its intended audience or, in the worst case, erroneous."

Lundgard wrote the paper with senior author Arvind Satyanarayan, an assistant professor of computer science who leads the Visualization Group in CSAIL. The research will be presented at the Institute of Electrical and Electronics Engineers Visualization Conference in October.

Evaluating visualizations

To develop the conceptual model, the researchers planned to begin by studying graphs featured by popular online publications such as FiveThirtyEight and NYTimes.com, but they ran into a problem—those charts mostly lacked any textual descriptions. So instead, they collected descriptions for these charts from graduate students in an MIT data visualization class and through an online survey, then grouped the captions into four categories.

Level 1 descriptions focus on the elements of the chart, such as its title, legend, and colors. Level 2 descriptions describe statistical content, like the minimum, maximum, or correlations. Level 3 descriptions cover perceptual interpretations of the data, like complex trends or clusters. Level 4 descriptions include subjective interpretations that go beyond the data and draw on the author's knowledge.

In a study with blind and sighted readers, the researchers presented visualizations with descriptions at different levels and asked participants

to rate how useful they were. While both groups agreed that level 1 content on its own was not very helpful, sighted readers gave level 4 content the highest marks while blind readers ranked that content among the least useful.

Survey results revealed that a majority of blind readers were emphatic that descriptions should not contain an author's editorialization, but rather stick to straight facts about the data. On the other hand, most sighted readers preferred a description that told a story about the data.

"For me, a surprising finding about the lack of utility for the highest-level content is that it ties very closely to feelings about agency and control as a disabled person. In our research, blind readers specifically didn't want the descriptions to tell them what to think about the data. They want the data to be accessible in a way that allows them to interpret it for themselves, and they want to have the agency to do that interpretation," Lundgard says.

A more inclusive future

This work could have implications as data scientists continue to develop and refine machine learning methods for autogenerating captions and alternative text.

"We are not able to do it yet, but it is not inconceivable to imagine that in the future we would be able to automate the creation of some of this higher-level content and build models that target level 2 or level 3 in our framework. And now we know what the research questions are. If we want to produce these automated captions, what should those captions say? We are able to be a bit more directed in our future research because we have these four levels," Satyanarayan says.

In the future, the four-level framework could also help researchers

develop machine learning models that can automatically suggest effective visualizations as part of the data analysis process, or models that can extract the most useful information from a chart.

This research could also inform future work in Satyanarayan's group that seeks to make interactive visualizations more accessible for blind readers who use a screen [reader](#) to access and interpret the information.

"The question of how to ensure that charts and graphs are accessible to screen reader users is both a socially important equity issue and a challenge that can advance the state-of-the-art in AI," says Meredith Ringel Morris, director and principal scientist of the People + AI Research team at Google Research, who was not involved with this study. "By introducing a framework for conceptualizing natural language descriptions of information graphics that is grounded in end-user needs, this work helps ensure that future AI researchers will focus their efforts on problems aligned with end-users' values."

Morris adds: "Rich natural-language descriptions of data graphics will not only expand access to critical information for people who are blind, but will also benefit a much wider audience as eyes-free interactions via smart speakers, chatbots, and other AI-powered agents become increasingly commonplace."

More information: Alan Lundgard et al, Accessible Visualization via Natural Language Descriptions: A Four-Level Model of Semantic Content, *IEEE Transactions on Visualization and Computer Graphics* (2021). [DOI: 10.1109/TVCG.2021.3114770](#)

This story is republished courtesy of MIT News ([web.mit.edu/newsoffice/](#)), a popular site that covers news about MIT research, innovation and teaching.