Online images may be turning back the clock on gender bias

February 14 2024, by Katie Gilbert

Credit: Solène Delecourt

A picture is worth a thousand words, as the saying goes, and research has shown that the human brain does indeed better retain information from

images than from text. These days, we are taking in more visual content than ever as we peruse picture-packed news sites and social media platforms.

And much of that visual content, according to new Berkeley Haas research published in the journal *Nature*, is reinforcing powerful gender stereotypes.

Through a series of experiments, observations, and the help of large language models, professors Douglas Guilbeault and Solène Delecourt found that female and male gender associations are more extreme among Google Images than within text from Google News. What's more, while the text is slightly more focused on men than women, this bias is over four times stronger in images.

"Most of the previous research about bias on the internet has been focused on text, but we now have Google Images, TikTok, YouTube, Instagram—all kinds of content based on modalities besides text," says Delecourt. "Our research suggests that the extent of bias online is much more widespread than previously shown."

Not only is online gender bias more prevalent in images than in text, the study revealed, but such bias is more psychologically potent in visual form. Strikingly, in one experiment, study participants who looked at gender-biased images—as opposed to those reading gender-biased text—demonstrated significantly stronger biases even three days later.

As online worlds grow more and more visual, it's important to understand the outsized potency of images, says Guilbeault, the lead author of the paper.

"We realized that this has implications for stereotypes—and no one had demonstrated that connection before," Guilbeault says. "Images are a

particularly sticky way for stereotypes to be communicated."

The extent of bias—and its effects

To zero in on gender bias in online images, Guilbeault and Delecourt teamed up with co-authors Tasker Hull from Psiphon, Inc., a software company that develops censorship-navigation tools; doctoral researcher Bhargav Srinivasa Desikan of Switzerland's École Polytechnique Fédérale de Lausanne (now at IPPR in London); Mark Chu from Columbia University; and Ethan Nadler from the University of Southern California. They designed a novel series of techniques to compare bias in images versus text and to investigate its psychological impact in both mediums.

First, the researchers pulled 3,495 social categories—which included occupations like "doctor" and "carpenter" as well as social roles like "friend" and "neighbor"—from Wordnet, a large database of related words and concepts.

To calculate the gender balance within each category of images, the researchers retrieved the top hundred Google images corresponding to each category and recruited people to classify each human face by gender.

Measuring gender bias in online texts was a trickier proposition—though one perfectly suited for fast-evolving large-language models, which noted the frequency of each social category's occurrence alongside references to gender in Google News text.

The researchers' analysis revealed that gender associations were more extreme among the images than within the text. There were also far more images focused on men than women.

Sticky images

The experimental phase of the study sought to illuminate the impacts that biases in online images have on internet users. The researchers asked 450 participants to use Google to search for apt descriptions of occupations relating to science, technology, and the arts.

One group used Google News to find and upload textual descriptions; another group used Google Images to find and upload pictures of occupations. (A control group was assigned the same task with neutral categories like "apple" and "guitar.")

After selecting their text- or image-based descriptions, the participants rated which gender they most associated with each occupation. Then they completed a test that asked them to sort various words into gender categories quickly. The test was administered again after three days.

The participants who worked with the images displayed much stronger gender associations compared to those in the text and control conditions—even three days later.

"This isn't only about the frequency of gender bias online," says Guilbeault. "Part of the story here is that there's something very sticky, very potent about images' representation of people that text just doesn't have."

Interestingly, when the researchers conducted their own online survey of public opinion—and when they looked at data on occupational gender distributions reported by the U.S. Bureau of Labor Statistics—they found that gender disparities were much less pronounced than those reflected in Google images.

Opening doors to new research

Delecourt and Guilbeault say they hope their findings lead to a more serious grappling with the challenges posed by embedded bias in online images. After all, it's relatively easy to tweak text to be as neutral as possible, whereas images of people inherently convey racial, gender, and other demographic information.

Guilbeault notes that other research has shown that gender biases in online text have decreased, but those findings may not reveal the whole story. "In images, we actually still see very prevalent widespread gender bias," he says. "That may be because we haven't really focused on images in terms of this movement towards gender equality. But it could also be because it's just harder to do that in images."

Guilbeault and Delecourt are already at work on another project in this vein to examine gender-age bias online using many of the same techniques. "Part of the reason this paper is so exciting is that it opens the door to many, many other types of research—into age or race, or into other modalities, like video," Delecourt says.

Provided by University of California - Berkeley