Fact-checking found to influence recommender algorithms

August 2 2023

Experiment Arms: — Fact−Checking — Fact−Checking + Voting
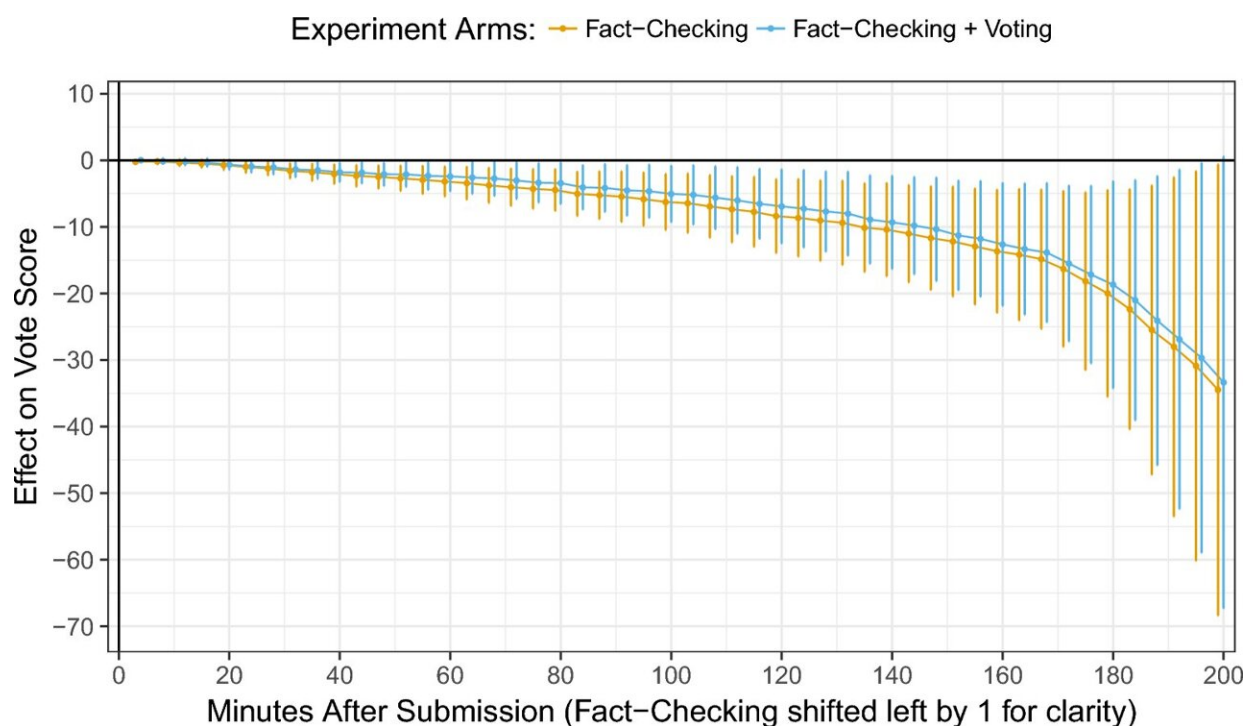
Both interventions reduced the vote score of the treatment groups compared to the control group. This chart shows the effect size and 95% confidence intervals for 50 linear regressions estimating the effect on vote score at a moment in time in the first 200 min after the article was posted. Credit: *Scientific Reports* (2023). DOI: 10.1038/s41598-023-38277-5

In January 2017, Reddit users read about an alleged case of terrorism in a Spanish supermarket. What they didn't know was that nearly every

detail of the stories, taken from several tabloid publications and amplified by Reddit's popularity algorithms, was false. Now, Cornell University research has shown that urging individuals to actively participate in the news they consume can reduce the spread of these kinds of falsehoods.

J. Nathan Matias, assistant professor of communication, conducted an experiment with a community of 14 million on Reddit and found that encouraging people to participate in knowledge-gathering could, in fact, move an algorithm's needle.

Suggesting that community members fact-check suspect stories, he found, led to those stories being dropped in Reddit's rankings.

"One of the lessons here is that we don't have to think of ourselves as captive to tech platforms and algorithms," said Matias, author of the study which published in *Scientific Reports*.

Working with the volunteer community leaders of the world news group on Reddit and focusing on news websites with a reputation for publishing inaccurate claims, researchers developed a software program that observed when a community member submitted a link for discussion. The software would then assign the discussion one of three conditions:

- Readers were shown a persistent message encouraging them to fact-check the article, and comment with links to further evidence refuting the article's claims.
- Readers were shown the message and encouraged to consider down-voting the article. Reddit articles are voted up or down by community members, known as "redditors," and ranked; accordingly.
- And in the control group, no action was taken.

Matias said he expected that fact-checking could backfire. If fact-checking focused more attention on an unreliable story, the algorithm might view it as positive reinforcement of the article and cause it to be ranked higher on average.

"There's the concern that if you repeat a falsehood often enough, is that going to anchor it in someone's mind?" he said. "And is there something similar with these recommendation algorithms that can't necessarily distinguish right from wrong? Even if you're fact-checking, the algorithm might see more engagement and show it to more people.'"

That turned out not to be the case. In a total of 1,104 news discussions, compiled from December 2016 to February 2017, Matias found that merely encouraging fact-checking—even without the down-voting prompt—caused a drop in story rank by an average of minus-25 spots. On Reddit, that would cause a story to drop off the front page, and likely be missed by a significant number of readers.

Though conducted on a narrow scale, Matias sees his experiment as proof that people can collectively take control of the information they are fed, and not just accept a steady diet of falsehoods.

"There's a lot of talk about this idea of determinism—that the decisions of an engineer somewhere in Silicon Valley influence our minds, and lock us into certain patterns of behavior," he said. "And while they do have influence, these systems are designed to react to humans. So, when people work together to improve our information environments, the algorithms can respond accordingly.

"This is a really powerful example of something that was very practical for this community," he said, "and is making fundamental contributions to scientific knowledge."