Unraveling Misinformation: Research on Social Media Correction Strategies

November 17, 2023

Editor's note: *Dr. Emily Vraga is an associate professor in the Hubbard School of Journalism and Mass Communication at the University of Minnesota. Her research focuses on how to detect and correct misinformation via social media, promote attention to higher quality news outlets online, and how to use news literacy messages to improve news consumption behaviors. Her research has appeared in journals such as Computers in Human Behavior, Science Communication, and Health Communication. She shares her research on mitigating the negative impacts of misinformation on social media as part of our Special Series on Addiction Myths and Misinformation.*

I didn't set out to study misinformation on social media. Early in graduate school, I became interested in how people experienced political disagreement. The rise of many prominent social media platforms – Facebook, Twitter, and YouTube especially – during this time meant I was increasingly thinking about this happened in these digital spaces. My colleague Leticia Bode sparked my interest in misinformation specifically in the summer of 2014. She and I had gone to graduate school together at the University of Wisconsin-Madison and both had our first academic appointments in the Washington, DC area. We met one summer day in a coffee shop in a grocery store in Virginia to talk about whether Facebook could use the new "related stories" functionality to respond when people share a misinformation link. It united our interests by thinking about the *quality* of the information people see on social media (Leticia's research) – and how whether disagreements were about opinions or *facts* (my research).

That project spurred a long line of research that I have done for the past decade testing different ways to mitigate the negative impacts of misinformation on social media. Most of my work in this space has been done with Leticia (when you find a good collaborator, you stick with them!). We like to use experiments, where we can precisely control the social media environments in which misinformation exposure happens and what corrections can look like. We've also tested these corrections across a lot of simulated platforms, including Facebook, Twitter, Instagram, and Facebook Live, and many other researchers have used similar approaches to study this question.

This work has a few major take-aways. First, we need to be thinking more about what makes correction effective not just in convincing the person sharing the misinformation – who is going to be the hardest person to persuade – of the best available evidence about the topic, but in reaching the much larger audience of people who are able to witness a public correction. We call this focus on the witnesses *observed correction* because it is those observers who are most likely to benefit from corrections. Our work and others have shown that corrections can help people become more accurate across a range of formats – telling people about the facts, explaining the misleading tactics that misinformation often relies upon to mislead, or telling personal stories about our experiences – and sources, like from expert organizations, algorithmic responses (like the Related Stories function in Facebook or a correction bot), or even from ordinary unknown social media users. This gives many people a lot of flexibility in responding to misinformation if they choose to do so.

Second, these benefits are theoretical; corrections can only work if many people are willing to perform them. Doing so can not only make people more accurate on the topic at hand but can hopefully reinforce social norms in favor of correction as something that *should* happen on social media. But we need to know a lot more about how these corrections happen on social media. Observational studies of social media shows that corrections do happen, but a lot of misinformation goes uncorrected. Leticia and I have recently interviewed over 60 people among the public and working at expert organizations devoted to high quality information about whether and why they publicly correct others. We found people told us about the many barriers to doing this kind of correction. Some of these – like concerns people won't like it or corrections won't work – are ones that we can address by talking about what research shows (people do like it and it does work). Other concerns are harder and will require changes to how platforms and publics

interact. Critically, this must include a commitment to reducing the toxicity and harassment that can plague online spaces where people disagree.

Lastly, I want to stress that correction is never going to be the only strategy for mitigating the problem of misinformation on social media; it will have to be combined with a lot of other efforts to really address the scope of the problem. We've called this a Swiss cheese approach to misinformation, because we need the various layers of protection that different efforts – like prebunking, news literacy, content moderation as well as correction – can provide to the vast problem of misinformation. Each of these layers has its holes (like slice of Swiss cheese) where misinformation can continue to spread, so we'll always need correction as the final layer to address misinformation.

— Emily Vraga, PhD

What do you think? Please use the comment link below to provide feedback on this article.