How do you solve a problem like misinformation?

Ryan Calo[1]*, Chris Coward[2], Emma S. Spiro[2], Kate Starbird[3], Jevin D. West[2]*

Understanding key distinctions between misinformation/disinformation, speech/action, and mistaken belief/conviction provides an opportunity to expand research and policy toward more constructive online communication.

The pandemic was planned. Climate change is a hoax. Joe Biden lost the election.

Trying to navigate misinformation about COVID, climate change, politics, and countless other topics can be overwhelming. This is true for the public, researchers, journalists, and policy-makers alike. As researchers dedicated to the study and resistance of misinformation, we often find ourselves in conversation with government officials and others trying to understand and address the phenomenon. To help illuminate the complexities of misinformation and to guide policy, we find three distinctions helpful: misinformation versus disinformation, speech versus action, and mistaken belief versus conviction (Fig. 1). Failing to appreciate these distinctions can lead to unproductive dead ends; understanding them is the first step toward recognizing misinformation and hopefully addressing it.

The first key distinction covers misinformation—erroneous or misleading information to which the public may be exposed, engage with, and share—and disinformation. Disinformation refers to a purposive strategy to induce false belief, channel behavior, or damage trust. Misinformation is usually discrete or standalone, as when a neighbor shares a false rumor or overhears a misleading exchange. Disinformation tends to take the form of a multifaceted campaign with a predetermined financial, political, or other objective. Disinformation campaigns blend orchestrated action and organic activity, relying on the participation of willing but unwitting online audiences.

The Plandemic video of May 2020, which called into question the origins of the pandemic, provides an interesting example (*1*).

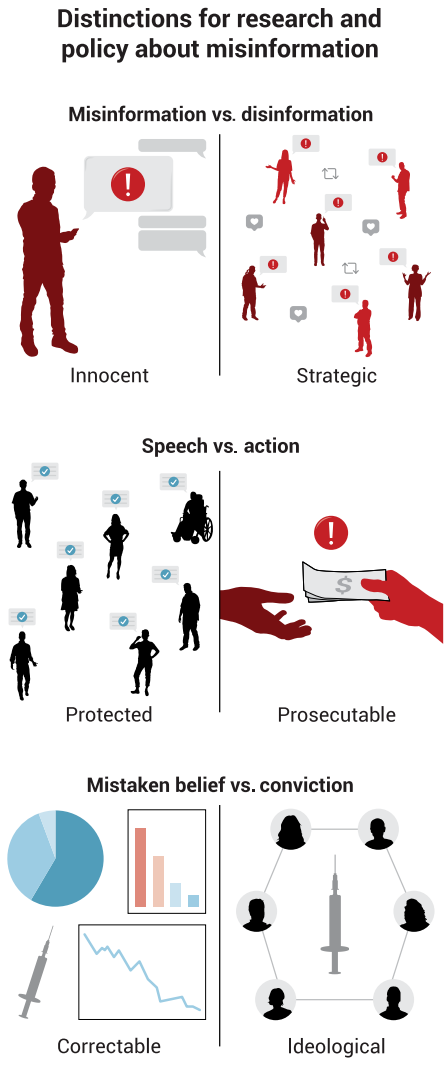

Fig. 1. Three distinctions for research and policy about misinformation. When addressing the issue of misinformation, three key distinctions exist to aid both policy and research: (i) misinformation versus disinformation, (ii) speech versus action, and (iii) mistaken belief versus conviction. Credit: Ashley Mastin/Science Advances.

The video featured a "whistleblower" with apparently sincere (if highly questionable) concerns about vaccination. Yet, the video also appears to figure into a broader disinformation campaign, i.e., a strategic effort to undermine vaccination efforts and to specifically attack the reputation of Anthony Fauci. There is evidence that a small group of online "influencers," motivated by political and ideological goals, coordinated to boost the whistleblower's online status and spread her message across multiple platforms before the video "going viral" (*2*).

Fighting misinformation is about identifying and addressing misleading messages. It is conceivable that a machine learning system could help flag misinformation or that legislation could define it. However, fighting disinformation is another matter. It is an exercise in disentangling the motivations of the various actors, some innocent and sincere, others strategic. The warning signs for a disinformation campaign may, ironically, involve true information and reasonable opinion. This suggests a need for researchers to follow people and strategies, rather than individual content alone, and for legislatures to address the problem at the level of incentives. The best way to address a foreign disinformation campaign, however, may be diplomacy and economic sanctions rather than artificial intelligence or tort law reforms.

A second distinction is the legal difference between speech and action. The U.S. Constitution protects free speech; however, it does not necessarily protect deceptive speech coupled with harmful action. This distinction potentially removes barriers to accountability for social media platforms that fail to address misinformation. Laws could require procedural safeguards and reporting about misinformation without censoring speech or treating Facebook or Google like a publisher. At a minimum, the assertive steps taken by technology companies to address coronavirus

[1]School of Law, University of Washington, Seattle, WA, USA. [2]Information School, University of Washington, Seattle, WA, USA. [3]Human Centered Design & Engineering, University of Washington, Seattle, WA, USA.
*Corresponding author. Email: rcalo@uw.edu (R.C.); jevinw@uw.edu (J.D.W.)

misinformation from warnings to outright deplatforming demonstrate potential methods to counter harmful content that has long plagued the internet.

We can and should ask more of internet platforms to address the conditions that they helped create and profit from, but what does the law say? The U.S. Constitution prohibits the government from censoring speech, even if the speech is misleading. Federal law (Section 230 of the Communications Decency Act), meanwhile, immunizes platforms like Facebook and Google from liability for speech on their platforms that originate from sources outside the company. Yet, neither the Constitution nor federal law grants legal protections for harmful conduct just because the action involves speech. For example, in striking down the Stolen Valor Act in 2012, which penalized lying about receipt of the Congressional Medal of Honor, the Supreme Court afforded states the power to criminalize fraud based on such a lie, and indeed, Congress passed a new version of the Stolen Valor Act the next year, with a requirement that lying about the medal had to be for the purpose of material gain for it to be criminal. In other words, the government can regulate doing things, or failing to do them, even if those things involve speech.

Consider again the Plandemic video. Different actors had different motivations for creating and boosting the video's messages. Were a person or organization to leverage the video to sell something harmful, to maliciously slander an individual, or otherwise knowingly deceive for material gain, a legal response may be tenable. Selling toothpaste that claims to cure COVID-19, for example, would be prosecutable. If the creators of the Plandemic video were simply misinformed or motivated by frustration or ideology, then their participation may not be prosecutable. In this case, the video's misleading claims may best be addressed through online interventions, such as authoritative banners on social media or other counter-speech efforts (*3*).

The final key distinction relates to the nature of belief itself, specifically, the difference between a mistaken belief and a conviction. We recognize that the distinction between belief and behavior is a subject of enduring interest in the social sciences. Indeed, one of our team's primary research questions examines how exposure to misinformation translates into both belief and behavior. Yet, the distinction between beliefs held out of mistake and beliefs held out of conviction remains undertheorized in both the research literature and within policy circles.

Vaccine hesitation offers a strong example of this distinction (*4*, *5*). Misinformation abounds, but we know that some people sincerely believe that vaccines are more harmful than helpful and oppose them on this basis. At the same time, it is possible that misinformation spread during the COVID-19 pandemic, like many of the claims in the Plandemic video, could recruit people that are not necessarily dogmatic in their views of vaccines initially but instead convinced by the falsehoods and persuasive storytelling.

In theory, if information about the COVID-19 vaccine's efficacy and safety is reported early and often through multiple channels or when conditions worsen, then some who are vaccine hesitant could be induced to change or reweigh their beliefs to accept a COVID-19 vaccine. Yet, for others, their opposition to vaccines may transcend scientific evidence. Some may hold or develop the conviction, as part of their cultural, social, or religious identity, that vaccines should be refused. This is why understanding conviction requires an awareness of the ideological frameworks that give rise to these convictions and hence necessitates engagement with psychologists, political scientists, and theologians.

Distinguishing between mistaken belief and convictions could help inform strategies about whether and how to correct misinformation. Researchers are beginning to recognize that the "backfire effect," the idea that corrections of misinformation could make things worse, is highly contextual, and evidence for the idea is mixed (*6*). Therefore, research investigating active corrections in online spaces is needed (*7*). As we develop and evaluate correction strategies, we will want to consider different approaches for correcting a false belief versus trying to change a potentially harmful conviction. Similarly, we want to consider upstream approaches that build greater resilience to misinformation before a problematic idea spreads.

Research into propaganda, conspiracy theory, and other distortions of information have a long history spanning multiple disciplines, but the rising tide of distorted and manipulative information has led to an increasingly visible, if still disparate, field of misinformation studies. This emerging field has the potential to inform our understanding of misinformation and the legal options to constrain misinformation and to advance our understanding of collective online behavior (*8*). Given the urgency and real-world impact of this issue, it is critical that the evolving policies in government and industry are informed by this research but also that the research itself is informed by these policy discussions. Our hope is that eventually, this collective work will make the field of misinformation studies obsolete.

REFERENCES

1. M. D. Kearney, S. C. Chiang, P. M. Massey, 'The Twitter origins and evolution of the COVID-19 "plandemic" conspiracy theory' (Harvard Kennedy School Misinformation Review, 1.3, 2020).
2. D. Alba, "Virus conspiracists elevate a new champion," *The New York Times*, 9 May 2020; www.nytimes.com/2020/05/09/technology/plandemic-judy-mikovitz-coronavirus-disinformation.html.
3. C. Geeng, T. Francisco, J. West, F. Roesner, Social media COVID-19 misinformation interventions viewed positively, but have limited impact. arXiv:2012.11055 [cs.CY] (2020).
4. A. Kata, Anti-vaccine activists, Web 2.0, and the postmodern paradigm—An overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine* **30**, 3778–3789 (2012).
5. E. Dubé, C. Laberge, M. Guay, P. Bramadat, R. Roy, J. Bettinger, Vaccine hesitancy: An overview. *Hum. Vaccin. Immunother.* **9**, 1763–1773 (2013).
6. T. Wood, E. Porter, The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Polit. Behav.* **41**, 135–163 (2019).
7. L. Bode, E. K. Vraga, See something, say something: Correction of global health misinformation on social media. *Health Commun.* **33**, 1131–1140 (2018).
8. J. B. Bak-Coleman, M. Alfano, W. Barfuss, C. T. Bergstrom, M. A. Centeno, I. D. Couzin, J. F. Donges, M. Galesic, A. S. Gersick, J. Jacquet, A. B. Kao, R. E. Moran, P. Romanczuk, D. I. Rubenstein, K. J. Tombak, J. J. Van Bavel, E. U. Weber, Stewardship of global collective behavior. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2025764118 (2021).

10.1126/sciadv.abn0481

ScienceAdvances

How do you solve a problem like misinformation?

Ryan CaloChris CowardEmma S. SpiroKate StarbirdJevin D. West

Sci. Adv., 7 (50), eabn0481. • DOI: 10.1126/sciadv.abn0481

Science Advances (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.