

INSIGHTS



POLICY FORUM

SOCIAL SCIENCE

The science of fake news

Addressing fake news requires a multidisciplinary effort

By David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, Jonathan L. Zittrain

The rise of fake news highlights the erosion of long-standing institutional bulwarks against misinformation in the internet age. Concern over the problem is global. However, much remains unknown regarding the vulnerabilities of individuals, institutions, and society to manipulations by malicious actors. A new system of safeguards is needed. Below, we discuss extant social and computer science research regarding belief in fake news

and the mechanisms by which it spreads. Fake news has a long history, but we focus on unanswered scientific questions raised by the proliferation of its most recent, politically oriented incarnation. Beyond selected references in the text, suggested further reading can be found in the supplementary materials.

WHAT IS FAKE NEWS?

We define “fake news” to be fabricated information that mimics news media content in form but not in organizational process or intent. Fake-news outlets, in turn, lack the news media’s editorial norms and processes for ensuring the accuracy and credibility of information. Fake news overlaps with other information disorders, such as misinformation (false or misleading information) and disinformation (false information that is purposely spread to deceive people).

Fake news has primarily drawn recent attention in a political context but it also has been documented in information promul-

gated about topics such as vaccination, nutrition, and stock values. It is particularly pernicious in that it is parasitic on standard news outlets, simultaneously benefiting from and undermining their credibility.

Some—notably First Draft and Facebook—favor the term “false news” because of the use of fake news as a political weapon (1). We have retained it because of its value as a scientific construct, and because its political salience draws attention to an important subject.

THE HISTORICAL SETTING

Journalistic norms of objectivity and balance arose as a backlash among journalists against the widespread use of propaganda in World War I (particularly their own role in propagating it) and the rise of corporate public relations in the 1920s. Local and national oligopolies created by the dominant 20th century technologies of information distribution (print and broadcast) sustained these norms. The internet has lowered the cost of entry to new competitors—many of which have rejected those norms—and undermined the business models of traditional news sources that had enjoyed high levels of public trust and credibility. General trust in the mass media collapsed to historic lows in 2016, especially on the political right, with

The list of author affiliations is provided in the supplementary materials. Email: d.lazer@northeastern.edu

51% of Democrats and 14% of Republicans expressing “a fair amount” or “a great deal” of trust in mass media as a news source (2).

The United States has undergone a parallel geo- and sociopolitical evolution. Geographic polarization of partisan preferences has dramatically increased over the past 40 years, reducing opportunities for cross-cutting political interaction. Homogeneous social networks, in turn, reduce tolerance for alternative views, amplify attitudinal polarization, boost the likelihood of accepting ideologically compatible news, and increase closure to new information. Dislike of the “other side” (affective polarization) has also risen. These trends have created a context in which fake news can attract a mass audience.

PREVALENCE AND IMPACT

How common is fake news, and what is its impact on individuals? There are surprisingly few scientific answers to these basic questions.

In evaluating the prevalence of fake news, we advocate focusing on the original sources—the publishers—rather than individual stories, because we view the defining element of fake news to be the intent and processes of the publisher. A focus on publishers also allows us to avoid the morass of trying to evaluate the accuracy of every single news story.

One study evaluating the dissemination of prominent fake news stories estimated that the average American encountered between one and three stories from known publishers of fake news during the month before the 2016 election (3). This likely is a conservative estimate because the study tracked only 156 fake news stories. Another study reported that false information on Twitter is typically retweeted by many more people, and far more rapidly, than true information, especially when the topic is politics (4). Facebook has estimated that manipulations by malicious actors accounted for less than one-tenth of 1% of civic content shared on the platform (5), although it has not presented details of its analysis.

By liking, sharing, and searching for information, social bots (automated accounts impersonating humans) can magnify the spread of fake news by orders of magnitude. By one recent estimate—that classified accounts based on observable features such as sharing behavior, number of ties, and linguistic features—between 9 and 15% of active Twitter accounts are bots (6). Facebook estimated that as many as 60 million bots (7) may be infesting its platform. They were responsible for a substantial portion of political content posted during the 2016 U.S. campaign, and some of the same bots were later used to attempt to influence the 2017 French election

(8). Bots are also deployed to manipulate algorithms used to predict potential engagement with content by a wider population. Indeed, a Facebook white paper reports widespread efforts to carry out this sort of manipulation during the 2016 U.S. election (5).

However, in the absence of methods to derive representative samples of bots and humans on a given platform, any point estimates of bot prevalence must be interpreted cautiously. Bot detection will always be a cat-and-mouse game in which a large, but unknown, number of humanlike bots may go undetected. Any success at detection, in turn, will inspire future countermeasures by bot producers. Identification of bots will therefore be a major ongoing research challenge.

We do know that, as with legitimate news, fake news stories have gone viral on social media. However, knowing how many individuals encountered or shared a piece of fake news is not the same as knowing how many people read or were affected by it. Evaluations of the medium-to-long-run impact on political behavior of exposure to fake news (for example, whether and how to vote) are essentially nonexistent in the literature. The impact might be small—evidence suggests that efforts by political campaigns to persuade individuals may have limited effects (9). However, mediation of much fake news via social media might accentuate its effect because of the implicit endorsement that comes with sharing. Beyond electoral impacts, what we know about the effects of media more generally suggests many potential pathways of influence, from increasing cynicism and apathy to encouraging extremism. There exists little evaluation of the impacts of fake news in these regards.

POTENTIAL INTERVENTIONS

What interventions might be effective at stemming the flow and influence of fake news? We identify two categories of interventions: (i) those aimed at empowering individuals to evaluate the fake news they encounter, and (ii) structural changes aimed at preventing exposure of individuals to fake news in the first instance.

Empowering individuals

There are many forms of fact checking, from websites that evaluate factual claims of news reports, such as PolitiFact and Snopes, to evaluations of news reports by credible news media, such as the *Washington Post* and the *Wall Street Journal*, to contextual information regarding content inserted by intermediaries, such as those used by Facebook.

Despite the apparent elegance of fact checking, the science supporting its efficacy is, at best, mixed. This may reflect broader tendencies in collective cognition, as well as

structural changes in our society. Individuals tend not to question the credibility of information unless it violates their preconceptions or they are incentivized to do so. Otherwise, they may accept information uncritically. People also tend to align their beliefs with the values of their community.

Research also further demonstrates that people prefer information that confirms their preexisting attitudes (selective exposure), view information consistent with their preexisting beliefs as more persuasive than dissonant information (confirmation bias), and are inclined to accept information that pleases them (desirability bias). Prior partisan and ideological beliefs might prevent acceptance of fact checking of a given fake news story.

Fact checking might even be counterproductive under certain circumstances. Research on fluency—the ease of information recall—and familiarity bias in politics shows that people tend to remember information, or how they feel about it, while forgetting the context within which they encountered it. Moreover, they are more likely to accept familiar information as true (10). There is thus a risk that repeating false information, even in a fact-checking context, may increase an individual’s likelihood of accepting it as true. The evidence on the effectiveness of claim repetition in fact checking is mixed (11).

Although experimental and survey research have confirmed that the perception of truth increases when misinformation is repeated, this may not occur if the misinformation is paired with a valid retraction. Some research suggests that repetition of the misinformation before its correction may even be beneficial. Further research is needed to reconcile these contradictions and determine the conditions under which fact-checking interventions are most effective.

Another, longer-run, approach seeks to improve individual evaluation of the quality of information sources through education. There has been a proliferation of efforts to inject training of critical-information skills into primary and secondary schools (12). However, it is uncertain whether such efforts improve assessments of information credibility or if any such effects will persist over time. An emphasis on fake news might also have the unintended consequence of reducing the perceived credibility of real-news outlets. There is a great need for rigorous program evaluation of different educational interventions.

Platform-based detection and intervention: Algorithms and bots

Internet platforms have become the most important enablers and primary conduits of fake news. It is inexpensive to create a web-

site that has the trappings of a professional news organization. It has also been easy to monetize content through online ads and social media dissemination. The internet not only provides a medium for publishing fake news but offers tools to actively promote dissemination.

About 47% of Americans overall report getting news from social media often or sometimes, with Facebook as, by far, the dominant source (13). Social media are key conduits for fake news sites (3). Indeed, Russia successfully manipulated all of the major platforms during the 2016 U.S. election, according to recent congressional testimony (7).

How might the internet and social media platforms help reduce the spread and impact of fake news? Google, Facebook, and Twitter are often mediators not only of our relationship with the news media but also with our friends and relatives. Generally, their business model relies on monetizing attention through advertising. They use complex statistical models to predict and maximize engagement with content (14). It should be possible to adjust those models to increase emphasis on quality information.

The platforms could provide consumers with signals of source quality that could be incorporated into the algorithmic rankings of content. They could minimize the personalization of political information relative to other types of content (reducing the creation of “echo chambers”). Functions that emphasize currently trending content could seek to exclude bot activity from measures of what is trending. More generally, the platforms could curb the automated spread of news content by bots and cyborgs (users who automatically share news from a set of sources, with or without reading them), although for the foreseeable future, bot producers will likely be able to design effective countermeasures.

The platforms have attempted each of these steps and others (5, 15). Facebook announced an intent to shift its algorithm to account for “quality” in its content curation process. Twitter announced that it blocked certain accounts linked to Russian misinformation and informed users exposed to those accounts that they may have been duped. However, the platforms have not provided enough detail for evaluation by the research community or subjected their findings to peer review, making them problematic for use by policy-makers or the general public.

We urge the platforms to collaborate with independent academics on evaluating the scope of the fake news issue and the design and effectiveness of interventions. There is little research focused on fake news and no

comprehensive data-collection system to provide a dynamic understanding of how pervasive systems of fake news provision are evolving. It is impossible to recreate the Google of 2010. Google itself could not do so even if it had the underlying code, because the patterns emerge from a complex interaction among code, content, and users. However, it is possible to record what the Google of 2018 is doing. More generally, researchers need to conduct a rigorous, ongoing audit of how the major platforms filter information.

There are challenges to scientific collaboration from the perspectives of industry and academia. Yet, there is an ethical and social responsibility, transcending market forces, for the platforms to contribute what data they uniquely can to a science of fake news.

The possible effectiveness of platform-based policies would point to either government regulation of the platforms or self-regulation. Direct government regulation of an area as sensitive as news carries its own risks, constitutional and otherwise. For instance, could regulators maintain (and, as important, be seen as maintaining) impartiality in defining, imposing, and enforcing any requirements? Generally, any direct intervention by government or the platforms that prevents users from seeing content

raises concerns about either government or corporate censorship.

An alternative to direct government regulation would be to enable tort lawsuits alleging, for example, defamation by those directly and concretely harmed by the spread of fake news. To the extent that an online platform assisted in the spreading of a mani-

festly false (but still persuasive) story, there might be avenues for liability consistent with existing constitutional law, which, in turn, would pressure platforms to intervene more regularly. In the U.S. context, however, a provision of the 1996 Communications Decency Act offers near-comprehensive immunity to platforms for false or otherwise actionable statements penned by others. Any change to this legal regime would raise thorny issues about the extent to which platform content (and content-curation decisions) should be subject to second-guessing by people alleging injury. The European “right to be forgotten” in search engines is testing these issues.

Structural interventions generally raise legitimate concerns about respecting private enterprise and human agency. But just as the media companies of the 20th century shaped the information to which individuals were exposed, the far-more-vast internet oligopolies are already shaping human experience on a global scale. The questions before us are how those immense powers are being—and

should be—exercised and how to hold these massive companies to account.

A FUTURE AGENDA

Our call is to promote interdisciplinary research to reduce the spread of fake news and to address the underlying pathologies it has revealed. Failures of the U.S. news media in the early 20th century led to the rise of journalistic norms and practices that, although imperfect, generally served us well by striving to provide objective, credible information. We must redesign our information ecosystem in the 21st century. This effort must be global in scope, as many countries, some of which have never developed a robust news ecosystem, face challenges around fake and real news that are more acute than in the United States. More broadly, we must answer a fundamental question: How can we create a news ecosystem and culture that values and promotes truth? ■

REFERENCES AND NOTES

1. C. Wardle, H. Derakhshan, “Information disorder: Toward an interdisciplinary framework for research and policy making” [Council of Europe policy report DGI(2017)09, Council of Europe, 2017]; <https://firstdraftnews.com/wp-content/uploads/2017/11/PREMS-162317-GBR-2018-Report-de%CC%81sinformation-1.pdf?x29719>.
2. A. Swift, Americans’ trust in mass media sinks to new low (Gallup, 2016); www.gallup.com/poll/195542/americans-trust-mass-media-sinks-new-low.aspx.
3. H. Allcott, M. Gentzkow, *J. Econ. Perspect.* **31**, 211 (2017).
4. S. Vosoughi et al., *Science* **359**, 1146 (2018).
5. J. Weedon et al., Information operations and Facebook (Facebook, 2017); <https://fbnewsroom.us.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf>.
6. O. Varol et al., in *Proceedings of the 11th AAAI Conference on Web and Social Media* (Association for the Advancement of Artificial Intelligence, Montreal, 2017), pp. 280–289.
7. Senate Judiciary Committee, Extremist content and Russian disinformation online: Working with tech to find solutions (Committee on the Judiciary, 2017); www.judiciary.senate.gov/meetings/extremist-content-and-russian-disinformation-online-working-with-tech-to-find-solutions.
8. E. Ferrara, *First Monday* **22**, 2017 (2017).
9. J. L. Kalla, D. E. Broockman, *Am. Polit. Sci. Rev.* **112**, 148 (2018).
10. B. Swire et al., *J. Exp. Psychol. Learn. Mem. Cogn.* **43**, 1948 (2017).
11. U. K. H. Ecker et al., *J. Appl. Res. Mem. Cogn.* **6**, 185 (2017).
12. C. Jones, Bill would help California schools teach about “fake news,” media literacy (EdSource, 2017); <https://edsources.org/2017/bill-would-help-california-schools-teach-about-fake-news-media-literacy/582363>.
13. Gottfried, E. Shearer, News use across social media platforms 2017, Pew Research Center, 7 September 2017; www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/.
14. E. Bakshy et al., *Science* **348**, 1130 (2015).
15. C. Crowell, Our approach to bots & misinformation, Twitter, 14 June 2017; https://blog.twitter.com/official/en_us/topics/company/2017/Our-Approach-Bots-Misinformation.html.

ACKNOWLEDGMENTS

We acknowledge support from the Shorenstein Center at the Harvard Kennedy School and the NULab for Texts, Maps, and Networks at Northeastern University. D.M.J.L. acknowledges support by the Economic and Social Research Council ES/N012283/1. D.M.J.L. and M.A.B. contributed equally to this article. Y.B. is on the advisory board of the Open Science Foundation. C.R.S. has consulted for Facebook. K.M.G. acknowledges support by the National Endowment for the Humanities.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6380/1094/suppl/DC1

10.1126/science.aao2998