

Studying human attention on the Internet

David Lazer^{a,1} 

“On the Internet, nobody knows you’re a dog” is the caption of a famous cartoon from the early years of the Internet. In the 1990s, this was meant as a light-hearted poke at the emerging medium. Today, this cartoon seems a prescient glimpse of a dystopian future, of an Internet of misleading Amazon reviews, fake news, and Russian trolls. “Falsehoods almost always beat out the truth on Twitter,” proclaimed *The Atlantic*, summarizing a major study in *Science* (1, 2). Yet, there has been relatively little science on the impact of inauthentic Internet content on what people believe and do. This gap reflects a major deficit in our collective capacity to conduct open science on the dynamics of human attention to content on the Internet (more below on this). Bail et al. (3) offer a critical evaluation of the potential impact of Russian troll content, finding little evidence that engagement with content in 2017 from (since revealed) Russian Twitter accounts is associated with increased polarization among US voters, one of the (inferred) objectives of Russian intervention in US political discourse.

The paper uses panel data of Twitter users coupled with 2 survey waves from 2017, collected for other purposes (4), with the following question: Is exposure to Russian troll content associated with political polarization (captured by an array of attitudinal and behavioral measures)? The answer—within these data—is robustly no.

This paper should therefore be reassuring that the effects of Russian social media interventions on polarization are minimal; however, readers should not sleep too easily at night. The paper, as with all papers, has important—and acknowledged—limitations. What is troubling, though, is that its core limitation, which is the lack of a strong measure of human attention, is almost universal to the research on human behavior on the Internet.

The Dearth of Data to Study Human Attention on the Internet

The Internet should be viewed as a global-scale cognitive psychology experiment, manipulating what people see and how they see it. The access that

science has to this information is quite limited, however. There is some research using browsing data (5)—collected by a small number of third-party vendors, who primarily sell data to industry—yet those data are rarely accessible to researchers and generally omit the massive amount of intraplatform exposure. There are some papers published by researchers from the platforms. The paper by Bakshy et al. (6) is exemplary, examining the content that people actually clicked on within Facebook; Vosoughi et al. (1) analyzed retweet rate contingent on exposure (what constitutes “exposure,” though, is not defined); and Eckles et al. (7) manipulated exposure to content to evaluate the causal impact of exposure on various behaviors. Given the sensitivity of platforms to negative news coverage that has since emerged in this space, however, it is unlikely that similar published research will be pursued in the future. The result is that when human attention is a key part of the asserted causal process (which it very often is) clumsy work-arounds are necessary.

As a result, there are remarkably few empirical field studies on exposure to content on social media, much less impact of exposure on cognition. There are few studies that provide simple statistics about online behavior, such as what the rate of retweeted content an individual is exposed to is, or even what proportion of all of the potential content in a timeline someone actually scrolls through. It is likely, however, that people are meaningfully exposed to just a tiny fraction of what could potentially appear in their timeline and only engage with a tiny fraction of what they are exposed to.

In Bail et al. (3), human attention is indeed a key step in the causal chain—presumably individuals will only be affected by the content that they actually saw and minimally cognitively processed. However, there is no easy way to evaluate which tweets even passed through people’s visual fields, much less what they attended to. The paper’s work-around is to capture what is trackable: what people retweeted, replied to, or liked, whom individuals followed, and whether trolls

^aNetwork Science Institute, Northeastern University, Boston, MA 02115

Author contributions: D.L. wrote the paper.

The author declares no competing interest.

Published under the [PNAS license](#).

See companion article on page 243.

¹Email: d.lazer@northeastern.edu.

First published December 17, 2019.

were mentioned by a followee. This is a sensible set of proxies, and the paper thoroughly evaluates the robustness of the findings to different constructions of the engagement variable. However, the gap between this and the reality of meaningful exposure might be quite large, and the fact that this is the best the field can do, in an area of such obvious scientific and public importance, is deeply troubling.

Further, engagement, by definition, is endogenous, which creates inferential challenges—for example, the paper shows that those who engaged with Russian content were individuals who were more interested in politics and had more ideologically homogeneous networks. The control group therefore includes individuals who may have been exposed and the most vulnerable to influence.

More generally, the lack of measures of exposure limits the development of the study of the use of the Internet that is informed by theories of human attention. It is plausible, for example, that the primary mechanism through which misinformation influences people is by repeated exposure. The rapid scanning by individuals of their timelines could lead to an illusory truth effect, where more familiar items are perceived as being more true (8). There is no conceivable way that the current state of the art in research in this space could evaluate such a process.

The objective here is not to critique the paper, which is appropriately circumspect and is an important contribution. A possible pathway of Russian influence was that it was stoking polarization via engagement with inflammatory content, and that seems significantly less plausible now. However, it is a major epistemological challenge for the field that other possible pathways of impact—of Russian influence and of misinformation and inflammatory content more generally—cannot be evaluated with the tools and data available to science today.

Building a Scientific Infrastructure for Studying Human Attention on the Internet

The lack of scientific infrastructure to study attention on the Internet is a solvable problem. There are two paths forward. In

principle, social media companies such as Twitter do have the ingredients for addressing these issues, or at least many more than those currently available to external researchers. Twitter does deserve credit for releasing the handles of the Russian accounts, which empowers research like this. However, Twitter knows far more: which content individuals might have seen or clicked on. Further, Twitter knows what types of exogenous changes to the platform might have affected whether individuals were exposed to given content, providing inferential leverage. However, while deep collaborations between academia and social media firms have been proposed (9), unsurprisingly there are few examples of such collaboration evaluating the possible negative impacts of online platforms on the quality of democratic discourse (for partial exceptions see refs. 1 and 10; the jury is still out on Facebook's efforts with Social Science One). Further, generally, access to social media data for research via application programming interfaces (APIs) has been declining, pointing to the need to develop models for social media research absent access to platform APIs (11).

It is possible that third-party audits of platforms such as Twitter will, at some point, be mandated, or that a model, such as Social Science One, will mature and effectively facilitate external (and possibly critical of the platforms) research. However, for the foreseeable future, the remaining alternative is for the relevant scientific community to build its own infrastructure, to capture what people are being exposed to, what the emergent processes of algorithmic curation are, and what the deeper effects on democratic discourse are. Ultimately, the only way to evaluate the risk from the things that go bump in the night is to turn on the light.

Acknowledgments

I appreciatively acknowledge comments from Nir Grinberg, Briony Swire-Thompson, Kenny Joseph, and Christopher Bail on earlier versions of the commentary. Errors belong to me alone.

- 1 S. Vosoughi, D. Roy, S. Aral, The spread of true and false news online. *Science* **359**, 1146–1151 (2018).
- 2 R. Meyer, The grim conclusions of the largest-ever study of fake news. *The Atlantic*, 8 March 2018. <https://www.theatlantic.com/technology/archive/2018/03/largest-study-ever-fake-news-mit-twitter/555104/>. Accessed 12 December 2019.
- 3 C. A. Bail et al., Assessing the Russian Internet Research Agency's impact on the political attitudes and behaviors of American Twitter users in late 2017. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 243–250 (2020).
- 4 C. A. Bail et al., Exposure to opposing views on social media can increase political polarization. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9216–9221 (2018).
- 5 A. Guess, B. Nyhan, J. Reifler, Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign (European Research Council, 2018) <http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf>. Accessed 12 December 2019.
- 6 E. Bakshy, S. Messing, L. A. Adamic, Political science. Exposure to ideologically diverse news and opinion on Facebook. *Science* **348**, 1130–1132 (2015).
- 7 D. Eckles, R. F. Kizilcec, E. Bakshy, Estimating peer effects in networks with peer encouragement designs. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7316–7322 (2016).
- 8 I. M. Begg, A. Anas, S. Farinacci, Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *J. Exp. Psychol. Gen.* **121**, 446–458 (1992).
- 9 S. Aral, D. Eckles, Protecting elections from social media manipulation. *Science* **365**, 858–861 (2019).
- 10 J. N. Matias, M. Mou, "CivilServant: Community-led experiments in platform governance" in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (ACM, New York, 2018), pp. 9:1–9:13.
- 11 D. Freelon, Computational research in the post-API age. *Polit. Commun.* **35**, 665–668 (2018).