


Ethical Considerations for Digitally Targeted Public Health Interventions

 See also Chou and Gaysynsky, p. S270.

Researchers, advocates, and policymakers increasingly worry that the Internet generally, and social media specifically, have become vectors of misinformation, manipulation, and other forms of malign influence.^{1,2} Unlike older forms of media, such as radio and television, Internet-driven influence differs in its capacity for individualized targeting, the speed with which messages can be transmitted and amplified, and the extent to which the creation and distribution of messages can be automated. While much attention has focused on the effects of such messaging on political discourse, researchers have traced equally concerning impacts on discussions pertaining to health-related issues, such as vaccine safety.³ Searching for ways to respond, public health officials and public health scholars have suggested a range of approaches, including increasing existing efforts to promote information and health literacy, devising strategies for publicly rebutting misinformation, and preparing clinicians and public health officials to address misinformation one on one.⁴

Such strategies are uncontroversial. Yet some contemplate going further, asking whether the same tools contributing to these problems—targeted, automated digital messaging—might be utilized to mitigate their negative

effects. For example, while acknowledging potential risks, Dunn et al. explore ways “social media data are used to predict or model health-related behaviours and outcomes” and “how these methods might be operationalised in the design of precision behavioural interventions.”⁵ One can imagine public health analogs of YouTube’s “redirect method,” which identifies users interested in terrorist or extremist videos and redirects them to antiextremist countermessaging.

Though research in this area is preliminary, it raises significant ethical questions that ought to be addressed in advance of further developments.

MANIPULATION AND AUTONOMY

In part, these proposals mirror ongoing debates about the use of so-called “nudging” to promote individual and public health, and they prompt some of the same normative considerations. Nudging involves shaping people’s choice environments in such a way that subtly steers them toward individually or socially beneficial decisions. Because such interventions are often designed to bypass people’s capacity for conscious deliberation, and function instead by triggering preconscious

decision-making heuristics (“cognitive biases”), they are fraught with questions about paternalism and manipulation.

While no consensus has been reached about the extent to which such worries are justified, they highlight morally relevant costs of intervening in people’s decision-making that might otherwise be neglected from the cost–benefit calculations public health officials have to make. We value autonomy (i.e., our capacity to make independent decisions), even when it means deciding to make ourselves worse off. Of course, situations can arise in which the potential harm is so grave that preventing it outweighs the cost of violating a person’s autonomy (e.g., in cases of suicidal ideation). But such situations ought to be treated as the exception rather than the rule. As researchers and public officials weigh the costs and benefits of utilizing digital influence strategies to promote health—especially precisely targeted (or “personalized”) interventions, which I and others

argue raise particularly acute manipulation worries²—they ought to seriously contemplate the costs of circumventing people’s capacity (and their right) to think and choose for themselves. If they decide to utilize such strategies, they should design interventions that targets can easily contextualize and understand—for example, by clearly indicating who is behind the messaging, why the person seeing it has been targeted, where they can find more information, and how they can opt out of future interventions.

PRIVACY

Privacy concerns arise because targeting individuals with relevant, timely public health messages requires collecting and processing information about them. One reason public health scholars are enthusiastic about the potential for these kinds of interventions is that ubiquitous digital technologies, such as smartphones and fitness trackers, create huge amounts of data that can be used to make predictions about individual and population-level health events. However, privacy scholars and advocates caution that the existence of such information does not entail that it

ABOUT THE AUTHOR

Daniel Susser is an assistant professor in the College of Information Sciences and Technology, a research associate in the Rock Ethics Institute, and an affiliate faculty member in the Philosophy Department at Penn State University, University Park, PA.

Correspondence should be sent to Professor Daniel Susser, College of Information Sciences and Technology, Penn State University, E325 Westgate Building, University Park, PA 16802 (e-mail: daniel.susser@psu.edu). Reprints can be ordered at <http://www.ajph.org> by clicking the “Reprints” link.

This editorial was accepted May 4, 2020.

<https://doi.org/10.2105/AJPH.2020.305758>

is “up for grabs.” People share information about themselves in particular contexts, with the expectation that it will be accessed by specific recipients and used for specified ends.⁶ Just as technology companies like Facebook and Cambridge Analytica faced backlash from the public when it learned they used information disclosed through social media to target political advertisements, public health officials ought to exercise caution before using such information to target health messaging, unless they have received clear, explicit, affirmative consent.

BIAS

Researchers have shown, time and again, that the algorithms used to deliver targeted content online are deeply susceptible to unintended, discriminatory bias. Using such tools to mitigate social media–driven misinformation or to promote truthful public health messaging thus raises the possibility of missing certain groups or targeting them with inaccurate information. As public health practitioners are unlikely to build message targeting systems themselves, relying instead on platforms like Google’s and Facebook’s, they ought to carefully consider the risk that their interventions might not reach all intended audiences (and indeed, that they might exclude already marginalized groups). At the very least, public health campaigns that do utilize ad targeting or other content recommendation platforms should be regularly audited to detect issues before they become widespread.

ACCOUNTABILITY

Finally, questions about accountability come to the fore whenever powerful institutions intervene in people’s lives. Such questions are especially urgent in this context because machine learning and artificial intelligence (the computational techniques that power most targeting and recommender systems) are known for their opacity, which derives from the fact that their inner workings are often protected by corporate trade secrecy laws, and their decision-making logics are difficult even for experts with proper access to understand.⁷ If public health organizations are going to use such tools ethically, they will need to go out of their way to create structures of transparency and accountability. That might involve storing messages for post hoc review, carefully logging who has seen them, and making that information readily available to auditors.

CONCLUSION

Dealing successfully with these ethical questions will require balancing difficult trade-offs. On one hand, the troves of personal data collected about each of us are incredibly revealing, and the tools for leveraging those data to target digital messages are powerful and readily available. It is easy to understand why researchers and public health practitioners are eager to explore the good they could do with them.

On the other hand, targeted digital public health interventions might also involve considerable ethical costs. The data that power targeting technologies are often collected in ways that disrespect data subjects’ privacy. Such technologies are liable to

target messages in ways that discriminate against marginalized groups. They create barriers to accountability. And targeted digital public health interventions threaten to influence our decision-making in ways that violate our autonomy. Whether the benefits of these interventions outweigh the costs should be determined on a case-by-case basis. To make such determinations, practitioners should consider both the severity of the health risks they are addressing (e.g., promoting healthy diets vs suicide intervention or combating health misinformation during a pandemic) and the extent to which they can minimize potential harms (e.g., whether messaging can be made transparent, and targeting data can be collected in ways that respect people’s privacy). Of course, the ethical issues discussed here are not exhaustive—rather, they suggest a place from which discussions about the ethics of targeted digital public health interventions can start. **AJPH**

Daniel Susser, PhD

ACKNOWLEDGMENTS

Thank you to the editors of this special issue for inviting me to comment, and to an anonymous reviewer for helpful feedback on an earlier draft.

CONFLICTS OF INTEREST

The author has no conflicts of interest to disclose.

REFERENCES

1. Marwick A, Lewis R. *Media Manipulation and Disinformation Online*. New York, NY: Data & Society Research Institute; 2017:106.
2. Susser D, Roessler B, Nissenbaum H. Online manipulation: hidden influences in a digital world. *Georgetown Law Technology Review*. 2019;4:1–45. <http://dx.doi.org/10.2139/ssrn.3306006>
3. Broniatowski DA, Jamison AM, Qi S, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *Am J Public Health*. 2018;108(10):1378–1384. <https://doi.org/10.2105/AJPH.2018.304567>

4. Chou WS, Oh A, Klein WMP. Addressing health-related misinformation on social media. *JAMA*. 2018;320(23):2417–2418. <https://doi.org/10.1001/jama.2018.16865>
5. Dunn AG, Mandl KD, Coiera E. Social media interventions for precision public health: promises and risks. *NPJ Digit Med*. 2018;1(1):47. <https://doi.org/10.1038/s41746-018-0054-0>
6. Nissenbaum H. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA: Stanford Law Books; 2010.
7. Burrell J. How the machine “thinks”: understanding opacity in machine learning algorithms. *Big Data Soc*. 2016;3(1). <https://doi.org/10.1177/2053951715622512>