OPINION

Opinion: Learning as we go: Lessons from the publication of Facebook's social-computing research

Jeffrey P. Kahn[a,1], Effy Vayena[b], and Anna C. Mastroianni[c]

[a]Berman Institute of Bioethics, The Johns Hopkins University, Baltimore, MD 21218; [b]Institute of Biomedical Ethics, University of Zurich, 8006 Zurich, Switzerland; and [c]University of Washington School of Law, Seattle, WA 98195

In the aftermath of the publication of the "emotional contagion" study conducted by Facebook researchers working with Cornell University scholars (1, 2), many observers weighed in about everything from the acceptability of undertaking the research at all, to how it was conducted, the rules and regulations that applied to it, and even the advisability of publishing the article. The views were disparate and conflicting (3–6). Our goal in this Opinion is not to try to settle these debates, but rather to attempt to draw some general lessons and offer recommendations from an ethics perspective as large-scale social-computing research moves forward. Our motivation is not to defend or chastise the research and technology communities or those responsible for ethics oversight of research. Instead, we wish to suggest that the development and application of an appropriate ethical framework and some form of ethics oversight is a moral imperative that is also in the interest of all. First, such oversight acts as a crucial signal of rules and accountability, which can increase overall trust and, in turn, the willingness to support and participate in research. Second, oversight leads to more credible research that others can build upon and funders and investors will support. Third, consistent approaches ease and encourage the trust and legitimacy needed for partnership and collaborations, which is the basis of 21st century science of all kinds.

The Facebook research example highlights areas of confusion and uncertainty with a growing category of investigation that is part innovation and part research, taking place in the context of very large user populations on Internet platforms that have become ingrained in the lives of users around the globe. We start with the presumption that large-scale social-computing research will continue, that it offers valuable insights, and that, like basic research in other scientific areas, its results are important to share. Advancing innovation is beneficial for companies, their users, and the public. Leading technology com-

panies consider this commitment to public benefit important (7–9), and it should be an important part of any discussion about relevant policies and practices. At the same time, questions have arisen about the need for research conducted in this environment to be subjected to some type of ethical framework and oversight to ensure that users' rights and interests are adequately safeguarded. Those benefits are unlikely to be fully realized, or at least to become accessible outside of private companies, without clarification and clear direction on a range of ethics-related issues.

Identifying Misfits and Gaps in Existing Approaches to Research Oversight

Large-scale social-computing research offers an environment that combines features of seemingly private behavior, public speech, social psychology research, and innovative technology development. This combination of features is a relatively recent phenomenon and a distinctive part of this fast-growing and evolving environment. Online environments and their privacy considerations were unimaginable at the time that foundational ethical frameworks for research were formalized and related regulations for protection of research subjects were developed in the late 1970s (10–12).

Those regulations, institutional policy and practices, and requirements for publication in research involving human participants are based on an understanding of research that does not account for these newer environments or for their scale. Federal regulations, for example, were designed to address inappropriate balancing of risks and benefits and questionable informed consent of subjects, and to assure voluntary participation of subjects, all informed by and in the context of worrisome exposure and scandals in biomedical research. Researchers from many nonbiomedical areas have long complained that applying clinical research rules makes little sense in the context of much social and behavioral research, and the interpretation

and application of regulations continue to evolve as a result (13, 14). As Fiske and Hauser recently argued in PNAS, research involving human participants in social-computing environments suffers from a similar mismatch of the realities of research and the policies governing it (15).

The inadequacies can be grouped into four categories: (*i*) the fit of existing regulation (what counts as research on human subjects, and oversight of private sector and collaborative research); (*ii*) requirements for and content of informed consent; (*iii*) confusion over the relevance of state and international borders; and (*iv*) clarification of criteria for research publication. These issues, separately and in combination, make it extremely difficult for even well-intentioned researchers to "do the right thing."

Regulatory Fit.
What counts as research with human participants? It is a consistent challenge for investigators and those responsible for the oversight of research to identify when data collection and analysis and related interventions should be treated formally as research with human participants under federal regulations. This determination is crucial for ascertaining whether under existing regulations institutional review board (or research ethics committee) review is required, which then becomes a condition for publication under publisher guidelines (16). We don't believe there is much controversy over whether the Facebook study involved "research," the traditional regulatory definition of which is ". . . a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge" (17) [45 CFR 46.102(d)], or whether it involved "human subjects," defined as ". . . a living individual about whom an investigator (whether professional or student) conducting research obtains (1) Data through intervention or interaction with the individual,

or (2) Identifiable private information" (18) [45 CFR 46.102(f)]. (As will be discussed below, there are reasons why the federal oversight process may still not apply to such research.) However, the questions raised during the ensuing debate suggest potential inadequacies in these definitions for many types of social-computing research (14). For example, do the definitions outlined above cover research on large social media platforms, such as Twitter and others? Some of the work being done in those settings, such as research on smoking-cessation programs, seems to be more like interventional research that fits within these definitions (19), whereas other research is collecting or analyzing "data" provided by millions of users speaking in social media's version of the public square, and may not necessarily fit within them.

Lesson #1: It is time for clear direction regarding what sorts of investigations in the context of social computing meet the definition of research on human subjects.

Ethics oversight of private sector and collaborative research. Social-computing research conducted or funded exclusively by private entities is not required to undergo the review and protections of United States federal research policies, although there may be other reasons—including publisher requirements—that companies choose to subject their research to oversight. If those companies partner or collaborate with institutions or researchers that are covered by federal regulations, it is possible that the research will be subject to the oversight policies applicable to that institution. When data are de-identified after collection for separate analysis by academic researchers, however, the data use no longer qualifies as "human subjects research" under regulatory definitions and may bypass research oversight, intentionally or accidentally. The gap in oversight leads to the odd possibility of identical research projects being performed under starkly different oversight practices, depending only on the source of funding and employment of the participating researchers.

Much of the discussion about the Facebook experiment focused on whether or not ethics review and approval was necessary. If we accept that research definitions are not as sharp and tidy as they once appeared, institutional review boards and research ethics committees need direction about when their approval is required in social-computing research contexts. The current oversight regime creates confusion in two respects. First, it forces artificial and arbitrary distinctions between research and nonresearch so the activity can be made to fit within the existing oversight scheme. Second, it trivializes the importance of ethical requirements by allowing the same activity to be held at different ethical standards. The nature of

social-computing environments and research using them requires ethical oversight that can address the entire continuum of the activity and not merely a portion of it.

Elsewhere, such as in several European countries, research regulation applies to all research activity independently of where it is conducted. However, two important features of this evolving research environment will make jurisdictional lines less relevant. First, the increasing number of public–private partnerships and collaborations involving data uses and reuses will raise challenging questions about balancing privacy and data sharing, as evidenced by the Facebook example and recent calls for large-scale data philanthropy projects (20). Second, there seems to be an increasing realization that the private sector has responsibilities to "respect, protect, and remedy" human rights, which will include a number of specific rights that are at stake in the research context (21). Whether avoiding research regulations is a strategic practice or a coincidence of private–academic research partnerships, it does not meet the intent of research policy to protect the rights and interests of research participants and to secure the social value of and trust in the research enterprise.

Lesson #2: Public–private collaboration is integral to large-scale social computing, and investigations should benefit from the relationship rather than be a source of confusion, inefficiency, and disincentive. Rules should consider research holistically rather than in piecemeal approaches resulting from artificial and outdated distinctions.

Informed Consent and Protection of Research Participants.
Research in social-computing environments shares some features of traditional social psychology research. It is intended to collect information about the behavior of individuals in environments under which conditions may be manipulated, often in efforts to explain why they behave the way they do. Social-computing research may have similar goals, making it a platform for not only social psychology but also other types of "traditional" research, or it may have goals more resembling product development, such as improving an existing Internet platform, technology, or service. Networking sites' terms of service often state that they conduct research. Users that have agreed to join such networks, therefore, technically consent to the possibility of their online activity being used in research. However, as we have argued elsewhere, this approach does not meet the requirements or the intent of informed consent in the sense that research ethics frameworks and guidelines intend (22). In an acknowledgment of this reality, the president of OkCupid noted candidly that in the terms

of service on their site, "you at least have the charade of consent . . ." (23).

Rather than debate whether informed consent as it has traditionally been practiced is being satisfied, the more relevant question is whether this concept of informed consent makes sense in social-computing research. Our view is that there are a number of reasons that its application in this setting is inapt and outdated. There are numerous parallel investigations (labeled as research or not) that involve uses of users' data, observations of online behavior, and even manipulation of the environment, as in some examples of social psychology research. Unlike in psychology research, however, participants in social-computing studies may not be recruited in the usual sense, and so may not even realize they are participating in research, let alone that there may be interventions, including manipulation or deception, involved. This seems to be the most sensitive aspect of the Facebook study: user-research participants were unaware of their participation, which included manipulation of their emotions. There cannot be traditional informed consent if this research is to go forward, and although some may argue that the answer is therefore to prohibit it, we believe that it can be performed ethically but requires an approach to disclosure and consent tailored to this environment.

Lesson #3: Approaches to informed consent must be reconceived for research in the social-computing environment, taking advantage of the technologies available and developing creative solutions that will empower users who participate in research, yield better results, and foster greater trust.

Confusion Over the Relevance of State and International Jurisdictions.
As a practical matter, worldwide access to social networks by a truly global pool of potential research participants means that geopolitical boundaries—state, regional, and international—mean much less than in the past, thereby complicating the determination of which laws govern in any particular setting and how to apply them. United States federal research policies currently address global research by requiring additional compliance with laws of the country where the research is being conducted, but it envisioned United States researchers conducting research or collaborating with sites in no more than a few countries at a time, rather than on the global scale of social computing. New approaches are required, involving at a minimum experts in areas outside of traditional research policy—such as Internet privacy—who are familiar with those legal challenges, to develop and potentially harmonize legal approaches to research

Kahn et al.

on a global scale. Such an attempt is currently under way in Europe with the revision of the European Data Protection Regulation, aspiring to harmonize data protection among European states (24).

Lesson #4: Although geopolitical boundaries may have limited impact on social-computing activities, confusion over jurisdiction and applicable laws remain an impediment to research. Harmonization will be a key to realizing the potential benefits of research on a global scale.

The Role of Publication Rules and Criteria. Scientific journals have a critical role to play in guiding the future of social-computing research. The International Committee of Medical Journal Editors' guidance for protection of research participants "when reporting experiments on people" created a common set of rules to be followed in order for research to be eligible for publication, independent of funding source, institutional affiliation, or country (25). With minor modifications of this existing guidance, journals have the opportunity to remove confusion and close the gaps in existing oversight. This should not be viewed as journal editors second-guessing local oversight and review, but as an opportunity to create common expectations of all published research, and to shed light on otherwise internal review within private companies. As the Facebook research example has indicated, there are roles and responsibilities for all those involved throughout the steps in research: investigators, companies, academic institutions, and the journal editors.

Lesson #5: As research areas evolve, journals can play an especially important role as the final and independent gatekeeper, assuring that research has been performed ethically.

Moving Forward

The upshot of the lessons outlined in this Opinion is that large-scale social-computing research would benefit from efforts to conceptualize an appropriate ethics framework to serve the many stakeholders involved, and fully realize the benefits and innovations that social-computing research has to offer. Effective approaches to any associated ethics

oversight will have to be nimble enough to adapt to unknown and unpredicted future research applications and sufficiently responsive to the time-sensitivities of private enterprise. It is in everyone's interest not to attempt forcing a 20th century regulatory regime onto 21st century technologies and approaches to research and innovation. Instead, it is necessary to craft a set of best-ethics practices that will serve as a common approach for all involved in large-scale social-computing research activities. Such a set of best practices will need to address and accommodate the context and distinctive features of social-computing research while drawing on fundamental ethical principles and approaches that have

guided international research policies over many decades. The broader research community, the private sector involved in social-computing activities, journal editors, regulators, and the users themselves are among the stakeholders that need to collectively engage in crafting this code of best practice.

The emotive reaction to the Facebook experiment is proof of the public interest in this set of issues, as well as an indication that best practices have yet to be identified. The future of social-computing research will evolve in ways that most of us cannot even fathom, let alone predict. The development of adaptable approaches to ensure that it is conducted ethically is critical to its success.

1 Kramer ADI, Guillory JE, Hancock JT (2014) Experimental evidence of massive-scale emotional contagion through social networks. *Proc Natl Acad Sci USA* 111(24):8788–8790.
2 Editorial Expression of Concern (2014) Correction to Kramer et al. 111 (24):8788. *Proc Natl Acad Sci USA* 111(29):10779.
3 Meyer MN (2014) Misjudgements will drive social trials underground. *Nature* 511(7509):265.
4 Watts DJ (2014) Lessons learned from the Facebook study. *The Chronicle of Higher Education Blog: The Conversation*. Available at http://chronicle.com/blogs/conversation/2014/07/09/lessons-learned-from-the-facebook-study. Accessed September 4, 2014.
5 Puschmann C (2014) All the world's a laboratory? On Facebook's emotional contagion experiment and user rights. Alexander Von Humboldt Institut Für Internet Und Gesellschaft. Available at www.hiig.de/en/all-the-worlds-a-laboratory-on-facebooks-emotional-contagion-experiment-and-user-rights. Accessed September 4, 2014.
6 Boesel WE (2014) Facebook's controversial experiment: Big tech is the new big pharma. *TIME.com*. Available at http://time.com/2951726/facebook-emotion-contagion-experiment. Accessed September 4, 2014.
7 Liu J, et al. (2010) Subjective sensing: Mission statement and a research agenda. Microsoft Research. Available at http://research.microsoft.com/apps/pubs/default.aspx?id=135844. Accessed September 4, 2014.
8 Page L, Brin S (2004) 2004 Founders' IPO letter. Google. Available at https://investor.google.com/corporate/2004/ipo-founders-letter.html. Accessed September 4, 2014.
9 Spohrer J (2013) Advancing big data for public good. *A Smarter Planet Blog*. Available at http://asmarterplanet.com/blog/2013/11/advancing-big-data-public-good.html. Accessed September 4, 2014.
10 Vayena E, Tasioulas J (2013) Adapting standards: Ethical oversight of participant-led health research. *PLoS Med* 10(3):e1001402.
11 Vayena E, Tasioulas J (2013) The ethics of participant-led biomedical research. *Nat Biotechnol* 31(9):786–787.
12 O'Connor D (2013) The apomediated world: Regulating research when social media has changed research. *J Law Med Ethics* 41(2):470–483.
13 National Research Council (2003) *Protecting Participants and Facilitating Social and Behavioral Sciences Research* (National Academies Press, Washington, DC).

14 National Research Council (2014) *Proposed Revisions to the Common Rule for the Protection of Human Subjects in the Behavioral and Social Sciences* (National Academies Press, Washington, DC).
15 Fiske ST, Hauser RM (2014) Protecting human research participants in the age of big data. *Proc Natl Acad Sci USA* 111:13675–13676.
16 International Committee of Medical Journal Editors (2013) *Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals*. Available at www.icmje.org/recommendations. Accessed September 4, 2014.
17 45. *Federal Register* 46.102(d) (2014).
18 45. *Federal Register* 46.102(f) (2014).
19 Prochaska JJ, Pechmann C, Kim R, Leonhardt JM (2012) Twitter=quitter? An analysis of Twitter quit smoking social networks. *Tob Control* 21(4):447–449.
20 Pawelke A, Tatevossian AR (2013) Data philanthropy: Where are we now? *United Nations Global Pulse Blog*. Available at www.unglobalpulse.org/data-philanthropy-where-are-we-now. Accessed September 4, 2014.
21 Ruggie JG (2011) Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework. United Nations Human Rights Council. 17th Session. March. (Office of the High Commissioner for Human Rights, New York and Geneva).
22 Vayena E, Mastroianni A, Kahn J (2013) Caught in the Web: Informed consent for online health research. *Sci Transl Med* 5(173):173fs6.
23 Cornish A (2014) OkCupid sometimes messes with a bit with love, in the name of science. *NPR.org*. Available at www.npr.org/2014/07/29/336356931/okcupid-sometimes-messes-a-bit-with-love-in-the-name-of-science. Accessed September 4, 2014.
24 Victor JM (2013) The EU general data protection regulation: Toward a property regime for protecting data privacy. *Yale Law J* 123(2):513–529.
25 International Committee of Medical Journal Editors (2014) *Protection of Research Participants*. Available at www.icmje.org/recommendations/browse/roles-and-responsibilities/protection-of-research-participants.html. Accessed September 4, 2014.