Topic Piece

Research Ethics
2016, Vol. 12(1) 35–43

DOI: 10.1177/1747016115579531
rea.sagepub.com

Facebook's emotional contagion study and the ethical problem of co-opted identity in mediated environments where users lack control

Evan Selinger
Department of Philosophy, Rochester Institute of Technology, USA

Woodrow Hartzog
Cumberland School of Law, Samford University, USA

Abstract

We argue a main but underappreciated reason why the Facebook emotional contagion experiment is ethically problematic is that it co-opted user data in a way that violated identity-based norms and exploited the vulnerability of those disclosing on social media who are unable to control how personal information is presented in this technologically mediated environment.

Keywords

big data, control, ethics, Facebook, identity, privacy

The problem of being co-opted

The Facebook emotional contagion experiment, in which researchers manipulated Facebook's news feed by, among other things, showing fewer positive posts to see

Corresponding author:
Evan Selinger, Department of Philosophy, Rochester Institute of Technology, 92 Lomb Memorial Drive, Rochester, NY 14623, USA.
Email: evan.selinger@rit.edu

if they would lead to greater user expressions of sadness, raises obvious as well as non-obvious problems (Kramer et al., 2014). Once reporters cycled through the obligatory "creepy as usual" stories about Facebook's emotion contagion experiment, research ethics became one of the dominant narratives of media coverage (Deterding, 2014; Grimmelmann, 2014a, 2014b). Typical inquiry revolved around the question of whether the procedures used to authorize research were adequate in light of the fact that hundreds of thousands of Facebook users were treated as human subjects. As a related matter, some attention was given to the issue of whether Facebook users actually could have been harmed by the experiment, given the specific stimuli involved.

In the immediate wave of popular media commentary, ethicists responded with mixed reactions. For example, Arthur Caplan (along with Charles Seife, 2014) concluded the case is best viewed as a "violation of the rights of research subjects." Similar dismay was expressed by Nicholas Evans, who stated that "the way the study was conducted was unethical" (Rosenbush, 2014a). Michelle N. Meyer (2014) argued that the situation serves as a powerful reminder that when academic scientists bound by the restrictions mandated by Institutional Review Boards (however imperfect) collaborate with private sector scientists, the potential for corporate power to be abused can be diminished. Patrick Lin pointed to the minimal transparency and expectations of privacy ordinarily found on Facebook, and expressed skepticism as to whether, given the typical state of affairs, the experiment qualifies as a lapse in ethical judgment (Rosenbush, 2014b).

However, the hyper-focus on the potential harm to those that viewed their personalized news feed has obscured an important ethical issue that implicates all who contribute to the social medium. By choosing to experimentally curate what information appeared on news feeds, Facebook appropriated user disclosures for the purpose of manipulating friends, family, co-workers, and the other types of people ordinarily connected by social networks. In doing so, Facebook *co-opted* users into a system that took information initially shared to meet socially laudable goals (e.g. stay in touch with loved ones) out of context, repurposed it, and deployed it—by way of secondary use—in ways that potentially could harm those we care about and try to behave responsibly towards (Selinger and Hartzog, 2014).[1]

At bottom, this aspect of the problem has two inter-related dimensions: 1) it highlights the limits of control users have over their disclosures in some mediated environments, and thereby the limited *agency* they can exhibit to ensure information is *responsibly disclosed*; and 2) it concerns companies imposing an *undesirable identity*—namely, collaborator—upon users as a cost of using information and communication services. In what follows, we elaborate on both points.

The limits of control over disclosures in mediated environments

'Control' is one of the most central concepts at stake in the debates over how to properly collect, use, and disclose personal information online. Indeed, much of US privacy law and a good deal of social expectations surrounding the proper use of information use revolve around control.

Daniel Solove, the John Marshal Harlan Research Professor of Law at the George Washington Law School, observes that privacy is regularly conceived of as an interest in controlling information—in determining who can access and correctly interpret information that directly or indirectly reveals something about us (Solove, 2008). Critics often argue that people who post their information online have only themselves to blame when things go awry because they can control what they type. This view focuses on the idea that people have it within their power to make better decisions, such as word things differently or refrain entirely from addressing sensitive issues.

This logic of control enables an overly reductive view of cyber-citizenship to gain legitimacy. Under this logic, a model internet user does his or her part by accepting responsibility for at least four general prohibitions: do not deliberately say something that hurts another person's feelings; do not say things about other people that they would not want disclosed; do not disclose information that can harm your own reputation; and use privacy settings to prevent prying eyes from peeking.

It is not hard to see why the ideal of control exerts a powerful hold on the imagination. Fixating on control suggests every time users log on to a site like Facebook they are responsible for doing the right thing and making good judgments because others can read what they write. As masters of our destiny, each of us appears to have the power to keep ourselves away from privacy harms and has a clear roadmap in sight for how to protect others.

Unfortunately, such an atomistic and choice-driven outlook ignores deeper structural realities and erroneously frames the common good as protected by each user exhibiting sensitivity and self-control. For as Solove argues, not only is the control paradigm less coherent and effective than regulators and judgmental peers typically presuppose, but it also unfairly places too much of the burden of "privacy self-management" on information disclosures themselves (Solove, 2013). In far too many situations, people lack the ability to make informed judgments about who, over time, will be able to access, share, and aggregate their information (Solove, 2013). They also have difficulty determining how much privacy can be given up and what harms can arise from disclosures that, at face value, appear innocuous and unrevealing (Solove, 2013). Additionally, the individualistic focus of managing one's own information with one's own private

good in mind makes it difficult for people to be sensitive to the adverse, "distributive effects" of disclosures that can be pernicious on a social level (Solove, 2013). Ultimately, due to cognitive limits, difficulties that mar how privacy policies are worded, corporate secrecy, and the accelerated pace of innovation that changes how information flows and can be analyzed by experts and technical systems, privacy scholars have expressed concern about the notice and consent regime being insufficient for ensuring that people can routinely make informed decisions when performing the gestures that signal agreement to an online company's Terms of Service (Hartzog, 2011).

One reason why it is especially hard to make informed choices when using products like Facebook is that we are not sure how our disclosures might be used by intermediaries with respect to other people. For example, Facebook can turn our online presence into other people's liabilities by studying our behavior and helping data brokers determine how others tick by virtue of demographic similarities. Indeed, when Facebook rolled out the interface changes that made it easier for users to enhance its mood graph by communicating through drop-down emoji, one of us lamented: "Even if you're someone who doesn't share anything, Facebook could potentially *reverse-engineer* your emotional persona by filling in the blanks from your like-minded friends' emotional states. In other words, the more your friends emote and translate their soulful moments into basic data points, the more Facebook can determine what makes you tick, too" (Selinger, 2013).

The problem of how our disclosures affect other people does not end there, with interdependence and similarities becoming marketing material. What the emotion contagion experiment brings to light is that privacy scholars and ethicists alike have failed to sufficiently appreciate just how little control users have in technologically mediated environments where they can only make decisions about what data they input into a medium but cannot determine what outputs the medium will produce for others when it translates their information into new forms and new contexts.

Imagine being a Facebook user who looked at the friendly "what's on your mind?" box and, with privacy settings engaged, typed in a banal but sad message about how poorly your day is going. That disclosure would appear to pass the control-based parameters of privacy management: nothing inappropriate is said about anyone else; nothing is said to people outside of one's circle; and nothing personal is said that is too intimate or risky. But if that information was entered when the emotion contagion experiment was running, the following outcomes could have arisen. Facebook could have altered what privacy scholar Helen Nissenbaum (2004) calls the expected "contextual integrity" of the information by re-purposing it as experimental stimuli that friends and family are presented with in order to adversely influence how they feel. All the while, you would be kept in

the dark—not knowing what your friends and family see when they log into Facebook and not knowing Facebook's motivation for showing them carefully curated information that you have contributed to—as the company avoids liability by using lengthy and obtuse language in a Terms of Service agreement (Facebook, 2014).

The bottom line is that while users can control the content of their disclosures in mediated environments, they cannot control how their disclosures are presented to others. Users cannot control the appearance and timing of their disclosures or any surrounding disclosures. This study sought to exploit that inability and demonstrate the significant impact the manipulation of presentation can have on the perceptions and effect of disclosures. This means that both disclosers *and* recipients of information are vulnerable to intermediaries. We believe our policy and norms should better reflect such power.

To enhance the clarity of our case and also make more explicit connections to the fundamental issues surrounding big data ethics, in the next section we will explain how the notions of control and responsibility at issue here are normatively salient because of their connection to matters that Neil Richards and Jonathan King (2013) have associated with identity.

The identity problem

Richards and King (2013) contend that a fundamental tension plagues our disclosures of personal information online. They call this problem the "identity paradox" (Richards and King, 2013: 43–4).

On the one hand, the personal statements we freely choose to make are integral parts of intersubjective processes that establish our cherished "identity". When we describe what we are doing, how we feel, what we think, etc., we create narrative threads that announce who we take ourselves to be and what we value. While we intend for these revelations to convey information to other people, we also take advantage of the archival nature of the internet and engage in social sharing to create digital memories our extended minds can revisit (Clark, 2003).

On the other hand, when we make personal disclosures on proprietary, for-profit platforms, the gestures also threaten our ability to be self-determining. Some companies know a lot about us by virtue of having accumulated significant portions of our digital dossiers. They also have in-depth knowledge of behavioral science and interface design, and provide services that customize the information users are presented with. Consequently big data giants can develop causally efficacious techniques to nudge us in directions they would like us to go: towards consumable goods, or, through filters, towards certain types of information (Richards and King, 2013: 44). Disclosing information online, therefore, can simultaneously bolster and jeopardize our ability to construct our identity.

Of course, one can only do so much to maintain control over personal identity. Luck can undermine our intentions, and others can form their own impressions of who we are and how we are behaving. In the end, the ability to manage the presentation of our identity to others is limited to casting probabilistic judgments about the likely outcomes that will follow from doing X or saying Y (Hartzog and Selinger, 2013). As Daniel Solove notes, "there are problems with viewing personal information as equivalent to any other commodity" because personal revelations are "often formed in relationships with others" who have minds and agendas of their own (Solove, 2008: 27).

In light of the basic constraints that prevent us from exhibiting absolute ownership of personal information, it would be unfair to hold Facebook—or any comparable online intermediary—accountable for the fact that other people misunderstand why we share certain information, respond to our posts in perverse ways, or flat out reject the terms we use to make sense of things. Yet deeming permissible these common uncertainties in establishing our identity is fundamentally different than sanctioning the creation of systems that undermine control by exploiting limited awareness and ignorance.

Until the details of the emotion contagion study were highly publicized, it seems likely that typical users were either: i) ignorant of the fact that Facebook curated their newsfeeds, or ii) erroneously believing that Facebook curation only aimed at presenting the most interesting and relevant stories (Gillespie, 2014). It seems implausible that the folks working at Facebook were unaware of this epistemologically deficient state of affairs—especially given how the company's business model is predicated upon ever more users disclosing increasing amounts of information.

To keep users comfortable with its service, Facebook does several things. For instance, it maintains community standards that protect against violence and threats, self-harm, bullying and harassment, hate speech, graphic content, and nudity. These civility norms are certainly laudable, but enforcement of these terms does more than just promote the social good. The prohibitions also are an instrumentally sound means of minimizing the type of toxic experiences that can lead unhappy customers to delete their accounts.

Another way Facebook keeps customers satisfied is by structuring user experience around positive interactions. Take the interface. It is designed to incline us to use the 'like' button and thereby constantly signal approval. It is not by dint of accidental omission that Facebook does not provide a dislike button for us to rapidly convey negativity (Gehl, 2013).

Finally, Facebook tries to keep users' morale high by minimizing the likelihood that when they use the service they will become anxious about the company's motivations or procedures. As with community standards, this means certain triggers need to be minimized. If users had a felt sense that they were being actively experimented on, some would quit. The same profit-impeding result would occur

if some users had a keen awareness that their virtuous identities were being converted into experimental collaborators whose behavior might harm others.

Think of it this way. When users interact with one another through Facebook's interface, they are only given highly selective information. Essentially they get to see what they post, some of what their friends share, and a non-overwhelming amount of advertising. Some of what remains invisible, therefore, is strategic. Crucially, users are not given windows into how Facebook's algorithms work (which is a trade secret) or how and why they get updated. Moreover, they only get the flimsiest peek into the scope of Facebook's advertising agenda.

In short, Facebook constructs a phenomenologically limited environment where what users see and experience does not prepare them to critically assess all of the relevant ways their identities are being distributed and reconstituted. While counterfactuals are hard to prove, it seems fair to say that if Facebook users were not deliberately kept in the dark, revelations of the emotional contagion experiment would not have proven explosive.

Conclusion

Even if Facebook's experiment resulted in a seemingly modest outcome and did not profoundly impact anyone's life, we need to keep in mind that a happy result couldn't be presupposed at the outset. If it could, there wouldn't be any need to run an experiment! Hypothetically, if a user shared a problem just to get it off her chest, it could have—combined with other attempts to do the same—been used in a way that made some of her friends (maybe ones with emotional disorders) sadder than it would have otherwise.

We have argued that the emotion contagion study exploited users' inability to determine how information is presented to others in the technologically mediated environment Facebook constructs. In so doing, it demonstrates the limits of relying on control as a central virtue of information ethics. As regulators, advocates, and consumers debate the best pathway forward for social media, we hope they take our analysis here into account and emphasize the moral harm that can occur when technology companies keep us in the dark about how our identities can be co-opted at the expense of others. It will take work, but protective measures can be integrated into many of the leading proposals, including contractual arrangements, like the People's Terms of Service (Selinger and Hartzog, 2014), Human-Subjects Research Oversight Committees (Calo, 2013; Salganik, 2014), and consent boxes that are explicit about users being willing study participants (Hill, 2014).

Acknowledgements

The authors would like to thank Emily Dreyfuss, editor of *Wired Opinion*, for initially soliciting our contribution to the debate over the Facebook controversy and supporting its expansion

into a more refined argument. We also appreciate Neil Richards and Jonathan King sharing their insights on how our respective articles overlap. And we benefitted from insights about how our thesis relates to issues of control shared by an anonymous reviewer.

Funding

This material is based upon work supported by the NSF under Grant No. 1140190. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF. The funding agency had no role in the writing of the manuscript or in the decision to submit the manuscript for publication.

Note

1. With permission of *Wired*, we are re-printing some of the prose initially found in Selinger and Hartzog (2014).

References

Calo R (2013) Consumer subject review boards: A thought experiment. *Stanford Law Review Online*. 66: 97–102.

Caplan A and Seife C (2014) Facebook experiment used Silicon Valley trickery. *NBC News*. Available at: http://www.nbcnews.com/health/mental-health/opinion-facebook-experiment-used-silicon-valley-trickery-n144386 (accessed 11 July 2014).

Clark A (2003) *Natural Born Cyborgs*. New York: Oxford University Press.

Deterding S (2014) Frame clashes, or: Why the Facebook emotion experiment stirs such mixed emotion. *Tumbling Conduct*. Available at: http://codingconduct.tumblr.com/post/90242838320/frame-clashes-or-why-the-facebook-emotion-experiment (accessed 11 July 2014).

Facebook Data Use Policy (2014) Available at: https://www.facebook.com/policy.php (accessed 19 November 2014).

Gehl R (2013) A history of like. *The New Inquiry*. Available at: http://thenewinquiry.com/essays/a-history-of-like/ (accessed 11 July 2014).

Gillespie T (2014) Facebook's algorithm—why our assumptions are wrong, and our concerns right. *Culture Digitally*. Available at: http://culturedigitally.org/2014/07/ facebooks-algorithm-why-our-assumptions-are-wrong-and-our-concerns-are-right/ (accessed 11 July 2014).

Grimmelmann J (2014a) As flies to wanton boys. *The Laboratorium*. Available at: http://laboratorium.net/archive/2014/06/28/as_flies_to_wanton_boys (accessed 16 July 2014).

Grimmelmann J (2014b) The Facebook emotional manipulation study: Sources. *The Laboratorium*. Available at: http://laboratorium.net/archive/2014/06/30/the_facebook_emotional_manipulation_study_source (accessed 16 July 2014).

Hartzog W (2011) Website design as contract. *American University Law Review* 60: 1635.

Hartzog W and Selinger E (2013) Big data in small hands. *Stanford Law Review Online* 66: 81–88.

Hill K (2014) Facebook manipulated 689,003 users' emotions for science. *Forbes*. Available at: http://www.forbes.com/sites/kashmirhill/2014/06/28/facebook-manipulated-689003-users-emotions-for-science/ (accessed 11 July 2014).

Kramer A, Guillory J and Hancock J (2014) Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings National Academy of Science* 111(24): 8788–8790.

Meyer MN (2014) Everything you need to know about Facebook's controversial emotion experiment. *Wired Opinion*. Available at: http://www.wired.com/2014/06/everything-you-need-to-know-about-facebooks-manipulative-experiment/ (accessed 11 July 2014).

Nissenbaum H (2004) Privacy as contextual integrity. *Washington Law Review* 79(1) 119–157.

Richards NM and King JH (2013) Three paradoxes of big data. *Stanford Law Review Online* 66: 41–46.

Rosenbush S (2014a) Facebook experiments tests ethics of manipulating emotion. *Wall Street Journal*. Available at: http://blogs.wsj.com/cio/2014/07/01/facebook-experiment-tests-ethics-of-manipulating-emotion/ (accessed 11 July 2014).

Rosenbush S (2014b) Facebook furor is overblown says Cal Poly ethicist. *Wall Street Journal*. Available at: http://blogs.wsj.com/cio/2014/06/30/facebook-furor-is-overblown-says-cal-poly-ethicist/ (accessed 11 July 2014).

Salganik M (2014) After the Facebook emotional contagion experiment: A proposal for a positive path forward. *Freedom to Tinker*. Available at: https://freedom-to-tinker.com/blog/mjs3/after-the-facebook-emotional-contagion-experiment-a-proposal-for-a-positive-path-forward/ (accessed 11 July 2014).

Selinger E (2013) The 'mood graph': How our emotions are taking over the web. *Wired Opinion*. Available at: http://www.wired.com/2013/08/move-over-social-graph-its-time-for-the-mood-graph-and-that-might-not-be-a-good-thing/ (accessed 11 July 2014).

Selinger E and Hartzog W (2014) How to stop Facebook from making us pawns in its corporate agenda. *Wired Opinion*. Available at: http://www.wired.com/2014/07/facebook-partner-in-scheme/ (accessed 11 July 2014).

Solove D (2008) *Understanding Privacy*. Cambridge, MA: Harvard University Press.

Solove D (2013) Privacy self-management and the consent dilemma. *Harvard Law Review* 126: 1880–1903.