

THE CHRONICLE OF HIGHER EDUCATION

RESEARCH

Big-Data Scientists Face Ethical Challenges After Facebook Study



Heather Ainsworth, The New York Times, Redux

Jeffrey Hancock, a Cornell U. professor who teamed up with Facebook on a controversial study of emotion online, says the experience has led him to think about how to continue such collaborations “in ways that users feel protected, that academics feel protected, and industry feels protected.”

By Paul Voosen | DECEMBER 15, 2014

Though it may not feel like it when you see the latest identity-affirming listicle shared by a friend on Facebook, we are a society moving toward evidence. Our world is ever more quantified, and with such data, flawed or not, the tools of science are more widely applied to our decisions. We can do more than observe our lives, the idea goes. We can experiment on them.

No group lives that ethos more than the life-hacking coders of Silicon Valley.

Trading on Internet-wired products that allow continuous updates and monitoring, programmers test their software while we use it, comparing one algorithmic tweak against another—the A/B test, as it’s known. As we browse the web, we are exposed to endless manipulations. Many are banal—what font gets you to click more?—and some are not.

Last summer the technologists discovered how unaware everyone else was of this new world. After Facebook, in collaboration with two academics, published a study showing how positive or negative language spreads among its users, a viral storm erupted. Facebook “controls emotions,” headlines yelled. Jeffrey T. Hancock, a Cornell University professor of communications and information science who collaborated with Facebook, drew harsh scrutiny. The study was the most shared

scientific article of the year on social media. Some critics called for a government investigation.

Much of the heat was fed by hype, mistakes, and underreporting. But the experiment also revealed problems for computational social science that remain unresolved. Several months after the study's publication, Mr. Hancock broke a media silence and told *The New York Times* that he would like to help the scientific world address those problems.

"How do we go about allowing these collaborations to continue," Mr. Hancock said more recently in an interview with *The Chronicle*, "in ways that users feel protected, that academics feel protected, and industry feels protected?"

Those problems will become only more acute as our quantified lives expand, Mr. Hancock said. "What's interesting with big data is that small, tiny effects that don't matter for the individual may matter at the aggregate," Mr. Hancock said.

There's wide agreement that the individual risks presented by the Facebook study were minimal. In effect, some subjects of the experiment saw minutely more posts containing positive or negative language, and subsequently, for every thousand words they shared, they posted one more matching emotional word. But nearly 700,000 people took part unwittingly in a psychology experiment. That scale presented a risk—perhaps not to each user but to science as a whole. "The harm here is to the reputation of science," said Christian Sandvig, an associate professor of information at the University of Michigan at Ann Arbor.

Mr. Hancock does not pretend to know the best way forward for big-data research. He knows the ethical qualms the study stirred up; he's faced them every day. "It led to a very difficult time for me personally," he said. "But it also was a difficult time for people who had concerns about this." Traveling to conferences since then, he has sought to be "somebody that people can bounce their ideas off of," he said, "because I went through it." And while some peers have criticized his continued presence, more seem curious about what he has to say.

"There are tough issues here," said Leslie Meltzer Henry, an associate professor of law at the University of Maryland's Carey School of Law and expert on research ethics. "Even the ethicists are divided."

Indeed, the emotion study has spawned a cottage industry of legal and ethical debate. Probably no institutional review board isn't pondering it. Few things are certain, but in talks with a variety of experts, one point is clear: When the whole population is up for experimentation, the experimental guidelines must change.

'Treat Your Users With Dignity'

For nearly a decade, Mr. Hancock, trained as a psychologist, has studied how increasing dependence on text in communication could alter how emotion is conveyed. Using Cornell students, he found that positive and negative emotion seemed to spread from one subject to another when they were interacting online. But those small experiments required a large intervention—having students watch a funny movie as they chatted, for example—for a very subtle effect.

At a conference several years ago, Mr. Hancock said, he met Adam D.I. Kramer, a data scientist at Facebook familiar with his work. For years, public concern had been rising that Facebook's users, by presenting rosy takes on their lives, could be making their friends sad in comparison. Fearing that the social-comparison effect was true, Facebook was going to test the reaction to a tweak in its ever-changing newsfeed algorithm, showing slightly fewer posts classified as containing positive or negative language. The same data, it turned out, could be used to hunt for evidence of emotional contagion. Without ethical review, over one week in early 2012, Facebook invisibly ran the trial. The rest is history.

So what should they have done differently?

Some critics of the experiment have argued for informed consent, a bedrock of federally supported research on human subjects; Facebook's users should have accepted the test, or at least opted into such trials. However, not all experiments with human subjects require informed consent: There are longstanding exceptions if the

experiment poses minimal risk and disclosure could bias results. Such exceptions are regularly used in social science, and it's likely the emotion study would have qualified for such a waiver had researchers asked for one, Ms. Henry said.

It's also quaint to think that users would click through the multiple dialogue boxes necessary to mimic informed consent, said Jonathan L. Zittrain, director of the Berkman Center for Internet and Society at Harvard University. Would you? Instead, he said, there ought to be independent proxies who represent the users and can perform that checking function.

"I worry about leaning too hard on choice," he said, "when the real thing is just treat your users with dignity."

Even if studies do not require consent, it could be standard to debrief subjects after the work is completed, Mr. Hancock said. (Facebook did not do that with the emotion study.) Send them to a page describing the study, and let them see as much information as they want. Few academics would seem to have a problem with that idea, though companies might be wary of scaring off users.

There's also the question of what experiments would require such disclosure. The line between design test and science experiment is fuzzy. The emotion study, if it hadn't been published, could be seen as Facebook running due diligence on a potentially negative user experience. Would that have made it less objectionable? Should only experiments done with academic collaborators face such scrutiny?

Whatever the costs, university scientists need to be held to a higher standard, said Mr. Sandvig. Despite scientists' own perceptions, society holds them in high regard—in much higher regard than Facebook, in fact. And if companies want to work in that world, "then it's a different set of standards," Mr. Sandvig said.

Facebook hasn't ignored the issue. In October it laid out a new research-review board and training for its engineers; it has also revised its terms of service to better reflect that it conducts research on users. While those guidelines are not perfect, if similar companies followed Facebook's approach, we'd be in a better place, Ms. Henry said.

Microsoft is developing similar guidelines, added Duncan J. Watts, a social scientist and principal researcher at Microsoft Research. If there's been one frustration for Mr. Watts in this post-emotion debate, it's been how few computational social scientists are speaking up.

The dominant participants in this conversation have been people who don't run online experiments, he said. They say that ethics is ethics, but Mr. Watts is not so certain. "If there's one thing I've learned," he said, "it's that ethical discussions in a vacuum are pretty meaningless."

The Creepiness Factor

Until Facebook's emotion study, social-media studies had been so focused on privacy protection that researchers missed other concerns. But it wasn't a violation of privacy that irked people.

"What bothered them was that it was creepy," Mr. Sandvig said. "That's not something you can argue them out of."

The creepiness factor isn't just about the idea of science run amok. When data scientists explain their practices, they don't often sound unreasonable, but negative connotations seem embedded in the language they use. "Experiments are bad," Mr. Watts said. "Manipulation is bad. Algorithms are cold and calculated and not human."

It could just be that society does not yet have a mental model for how this machine-mediated online world works. Several researchers compared the dawning awareness that came with the emotion study to debates about advertising in the mid-20th century, when people discovered subliminal messaging.

"Now we have a good idea how ads work," Mr. Hancock said. "We're not going to get upset seeing ads manipulating emotions."

Mr. Hancock has begun new work spurred by that idea, looking at the models people use to think about algorithms. He's keeping the tests in the lab for now. "Once I have

a better understanding," he said, "I'd like to expand the scope of it."

Paul Voosen is a senior reporter covering the sciences. Write him at paul.voosen@chronicle.com; follow him on Twitter @vooos; or see past work at voosen.me.

Copyright © 2018 The Chronicle of Higher Education