

We Are All Research Subjects Now

And Cold War-era safeguards won't protect us

By Sarah E. Igo | OCTOBER 07, 2018



Oliver Munday for The Chronicle Review

This spring, with some fanfare, Facebook and the Social Science Research Council (SSRC) announced that they would team up for a novel research collaboration. The unusual partnership has been greeted with equal parts praise and criticism. What seems undeniable, however, is what the project represents: the first outlines of a 21st-century social-research complex shaped more by big data than conventional data sets, and by corporate rather than public backing. Given its evident importance, those leading the effort will need to resist a reflexive retreat to old frameworks for ethical inquiry even as they venture into new territory.

While many of the specifics are still being hammered out, the joint SSRC-Facebook Social Data Initiative marks the first time that the social-media giant has agreed to release large amounts of proprietary data to outside scholars. Under the umbrella of the initiative's "Social Media and Democracy" program, the recipients will in the first instance be academics analyzing issues of immediate import: the impact of social media on elections and politics around the globe.

Social scientists have long sought access to Facebook's informational riches, but with limited success. Although the company has come under fire for its looseness in sharing user data, that same laxity has never applied to credentialed scholars. But recent events

have provided an opening. Facebook is under duress, reeling from the Cambridge Analytica scandal, a series of related federal and international investigations, heightened public scrutiny of its privacy practices, and a nose dive in the value of its stock.

For its part, the SSRC notes that its organizational mission — analyzing "social, cultural, economic, and political processes" — requires reckoning with the fact that such processes increasingly take shape online. Few scholars would deny the allure of Facebook's enormous trove of social and behavioral data, even if those data present special challenges. For those who wish to understand peer influence, media effects, and political manipulation, Facebook possesses information about more than a billion users — a data set with a breadth and depth unlike any other.

The stumbling block has been scholars' lack of independent access to closely guarded, proprietary corporate material. The SSRC's carefully brokered Facebook partnership, funded by seven nonprofit foundations, is intended to solve that problem. As one official statement has it, "realizing the potential of social science under these new conditions requires tripartite cooperation among academics, the private sector, and government." The Social Data Initiative will test the possibility that corporate data gathering and mining, designed to maximize profits, can be turned to socially useful ends. Establishing "a new paradigm for research collaboration between industry and the academy," the SSRC avows, "can produce findings that improve everybody's lives."

And yet it will surprise no one familiar with the history of postwar U.S. academic research — books like Mark Solovey's *Shaky Foundations* (Rutgers University Press, 2013) or Joy Rohde's *Armed With Expertise* (Cornell University Press, 2013) — that this arrangement sets off alarm bells. The venture invites comparisons to earlier instances of heightened attention to "sponsored" research, most notably the entanglement of the national-security state and the academy during the Cold War. Revelations in 1965 about the military's Project Camelot, which tapped social scientists in order to understand and influence the internal affairs of "target" countries, for example, led to much soul-searching about knowledge conscripted — and corrupted — by the priorities of the national-security state. Ever since, the question of patronage and its price has troubled the ideal of academic objectivity.

Science-studies scholars, of course, have long understood that research can never be "pure" in the sense that it is untouched by cultural assumptions and biases, or by the realities and politics of its funding. Still, the fact that key threats to human and behavioral research now flow from a social-media company rather than the military or the government has triggered fresh concerns about scholarly access and independence. In an echo of debates from the 1960s concerning the "military-industrial-university complex," today's skeptics ask how the scientific integrity of an operation can be assured when the profits and reputation of a corporation as powerful as Facebook are at stake.

In response, Facebook has pledged that its role in the research process will end once it has handed over its data sets, and the arrangement stipulates that the company will exercise no pre-publication approval. Likewise, the collaboration's funders will have no access to the data. Furthermore, an expert commission appointed by the SSRC will oversee the competitive grant process whereby scholars apply to make use of the company's data, and all proposals will be vetted by peer review.

It is possible that these provisions, drafted partly in response to past debacles involving sponsored research, will be enough to satisfy critics. Yet the question of scientific standards is only the first of the hurdles to the collaboration that the SSRC will have to clear. The second, and likely more vexing one, concerns protections for human subjects.

Here too, the Social Data Initiative is inevitably yoked to debates from the 1960s about research subjects' rights to dignity and privacy. For better or worse, the SSRC relies on scholarly regulations of the same vintage. Noting that data sets will be anonymized by Facebook, the organization promises that all research projects vying for those data "must first have been reviewed by a Common Rule-compliant University Institutional Review Board (IRB), federally approved IRB, or international equivalent," as well as a data-ethics committee, and will be "bound by the data access and privacy laws where they operate."

In so doing, the SSRC makes recourse to a familiar scholarly apparatus. But this standard of protection raises as many questions as it answers. How will data that index a wealth of personal information be securely anonymized in an age of sophisticated re-identification

techniques? And how will the rights of upwards of a billion Facebook users be preserved — that is, guided by the rules of institutional review and the principle of informed consent — once enmeshed in this new research complex?

Indeed, we might ask whether the solutions of another century and a markedly different research context can be applied to a big data project of the sort that the SSRC-Facebook collaborators envision. Revisiting the circumstances that gave rise to modern human-subjects regulations may hold some answers.

In 1966, a fledgling sociology Ph.D. candidate at Washington University in St. Louis planted himself in a public restroom, prepared to embark on his dissertation research. To those in the know, the site was a "tearoom," a reliable venue for outwardly straight men to engage in quick and surreptitious homosexual encounters. And so Laud Humphreys was there, too, in pursuit of social-scientific insights into the coordination and codes of covert public sex. Fashioning himself as an insider to gain access to the tearooms, Humphreys observed men in the midst of fellatio, charted their dialogue and movements, and even took down their license-plate information so that he could later interview them under another pretext.

This was — in the eyes of Humphreys and his advisers — trailblazing research. Hardly anyone in the field had investigated homosexual behavior at such close range. His account, published in 1970 as *Tearoom Trade: Impersonal Sex in Public Places*, garnered the prestigious C. Wright Mills Award from the Society for the Study of Social Problems.

But Humphreys published his research at a moment of turmoil in the social sciences, with parallels to our own moment of replication challenges, corporate-bankrolled research, and popular skepticism of academic expertise. Broadly speaking, the 1960s crisis was about challenges to the unchecked authority of scholars over knowledge creation. It was propelled only in part by exposés of military-sponsored research during the Cold War. It was fueled as well by a number of well-publicized scandals that laid bare the human costs of social investigation.

While many of these concerned the biomedical sphere — with one flash point Henry Beecher's 1966 exposé of disturbing experiments conducted on "mentally defective" children, the elderly, charity patients, alcoholics, and the terminally ill — academic social scientists were hardly immune. Leading the pack was the social psychologist Stanley Milgram. His "obedience to authority" experiments in the early 1960s, which led subjects to believe that they had willingly applied painful electric shocks to confederates, provoked outrage. More indignation followed when the psychologist Philip Zimbardo announced that he had terminated his Stanford Prison Experiment early to stanch the cruelty that undergraduates acting as guards inflicted on fictive prisoners.

Charges of exploitation and damage in experimental settings prompted new calls for research subjects' rights in the later '60s and early '70s. Worries about medical and social-scientific abuses, first raised amid revelations of Nazi practices during World War II, would in this period finally force a public debate on research ethics within the United States.

More generally, assertions about the autonomy, dignity, and privacy of research subjects reflected public unease with the sheer invasiveness of the era's social inquiry — what the University of Chicago sociologist Edward Shils characterized as the "diminution of inhibitions on intrusiveness into other persons' affairs." Conveners of a social-science task force in the 1970s noted a "rapidly increasing concern about the protection of the rights and welfare of human subjects." They ticked off six possible risks in social research: coercion, deception, invasion of privacy, breach of confidentiality, stress, and collective damage. The harms ran the gamut from "psychological injury" to "loss of self-esteem" and "generation of self-doubt." But the feature that most distinguished social research, they contended, was its special capacity to invade privacy.

Indeed, even as he filed his dissertation, Laud Humphreys faced an ethics inquiry from the university administration, triggered by the new human-subjects review rules attached to a federal grant that had funded his research. Following the publication of *Tearoom Trade*, a raft of academic critics and journalists charged that the study — no matter how valuable its sociological findings — had crossed a line.

Humphreys, his defenders pointed out, had taken every precaution to protect the confidentiality of his subjects. He did not record names or other identifying information on his questionnaires and kept the master list in a safe-deposit box a thousand miles away. These measures were state-of-the-art research ethics in 1966. Nevertheless, many charged that Humphreys had trampled the dignity and the rights of his subjects. By invading the privacy of unwitting tearoom users, he had opened them to both physical and psychological risks.

Most problematic, because Humphreys had relied on deception to observe his subjects, he had never even made them aware of those risks. Social scientists, argued a fellow sociologist, "assume a great responsibility" — indeed, too great a responsibility — "when they deny respondents the rights of voluntary participation." He said that "those whom we hope to help by our efforts" ought "not be put in such jeopardy except with their explicit consent."

For their part, Laud Humphreys's mentors resolutely defended the primacy of the researcher's "right to know." They pinned criticism of *Tearoom Trade* to a reaction against social inquiry itself. Not surprisingly, Humphreys agreed that "the pursuit of truth, the creation of countervailing knowledge, the demystification of shadowy areas of human experience" were all worth the possible risks to privacy. For him, the chief ethical question in the tearoom study was the relatively straightforward one of safeguarding his subjects' confidentiality so that they could not be identified by outsiders.

But it was his critics, and not Humphreys, who would carry the day. Whatever other scholarly contributions it would make, *Tearoom Trade*'s most lasting legacy was to solidify a new ethical matrix for the researcher-subject relationship. By 1979 the tearoom study could be characterized in the pages of *Science*, alongside Stanley Milgram's and Philip Zimbardo's experiments, as "a classic in the fast-growing field of ethics in social science research, where it is commonly cited as a crass violation of subjects' rights." The same article explained that, no matter how protective of the subjects' confidentiality, a project like Humphreys's "is regarded as indefensible in the ethical climate of the late 1970s."

Even before then, the study had become a common reference point for social scientists and discussions of research ethics. The sociologist's foray into the tearrooms influenced a series of efforts attempting to strike the proper balance between a right to research and the rights of the researched: the trade-off, as one key scholarly commission framed it, between a "gain in knowledge" and the "cost in privacy."

While these reports often shied away from explicitly recognizing the rights of human subjects in favor of professional discretion and self-policing, they affirmed the responsibility of the researcher to foster "conditions that give fullest protection to individual human dignity." The outcome was an array of new agencies, regulations, and policies, including the 1974 National Research Act, which created the National Commission for the Protection of Human Subjects in both biomedical and behavioral research. The new standards stipulated informed consent and prior review, typically under the mantle of institutional review boards.

Fittingly, given his part in creating it, even Laud Humphreys eventually came around — albeit reluctantly — to the new research regime. In line with freshly drafted regulations on human subjects, he advised his own students to obtain the consent of all research participants before interviewing them. And he stated that were he to begin the tearroom study all over again, he would spend an additional year cultivating willing respondents rather than observing them unawares, no matter what information would be lost.

These brakes on social inquiry are the same ones that academic researchers labor under — and sometime chafe under — today. And they are the same ones that the Social Data Initiative enlists in its public statements. But we should note that they were designed for research situations like the tearroom ethnography: where the privacy intrusion was intentional, where the potential harm to individuals' dignity was obvious, where specific consent from the human subjects might feasibly have been obtained, and where careful scholarly review might have prompted a more ethical research design.

The data sets that Facebook plans to hand over to SSRC-approved researchers are by nature quite different. They first of all are being used only after the fact, having been collected via no peer-review process by a for-profit company. Accepting the terms of

service of a social-media company is a far lower bar than the "informed consent" required by an IRB. These will be reams of personal data, possibly quite sensitive, and gathered unobtrusively, without the express consent (and often, knowledge) of those being researched.

It is not even certain whether the donors of data in this new venture are "research participants" in the sense that social scientists of the last century would have recognized. Some commentators have argued that because the company's data will be anonymized before researchers get ahold of them (itself a concern as re-identification techniques improve), standards of informed consent do not even pertain.

Can the protections intended for a relatively small group of identifiable subjects in a bounded study — the men in St. Louis's public restrooms in 1966, say — be extrapolated to the more than one billion virtual subjects who have been swept willy nilly into Facebook's informational cache? The SSRC, in its early statements about the Social Data Initiative, seems to believe so. But today's system of IRBs and federal regulations ought not be treated as definitive, especially given the new risks and possibilities presented by industry partnerships and big data.

Rather than accept the solutions of the 1960s and 1970s as a given, the new initiative would do better to reopen the questions that *Tearoom Trade* and other cutting-edge social research of its day generated about the legitimate bounds of social inquiry. The regulations that emerged were important, but so was the larger claim that human dignity ought to serve as an essential check on research ambitions.

Even then, debates over subjects' rights recognized that IRBs and informed consent were imperfect mechanisms. They could not resolve every ethical quandary. Confidentiality, even in the hands of scrupulous researchers, could be violated. Consent — if arrived at through coercion, even subtle forms of pressure, or incomplete information — could be fictive.

Such critiques were aired a half-century ago. But they are yet more urgent, and confounding, in our big-data present, as a new raft of corporate scandals, ranging from Equifax to Cambridge Analytica, makes clear. High-level data breaches, silent tracking of

all manner of online behavior, and innovations in the techniques of psychological suasion have raised the stakes on both confidentiality and consent. Some, indeed, question whether those principles are attainable in the social-media realm.

For those who care both about pathbreaking social research and the rights of human subjects, the SSRC-Facebook collaboration poses dilemmas equivalent to those raised by *Tearoom Trade*. It is an opportunity to reconsider, and possibly revise, the rules of social inquiry. Are the guidelines for ethical research and treatment of subjects that were devised nearly 50 years ago a durable resource for us today? What kind of help can these tools, forged in quite different conditions, offer us in resolving the potential privacy violations and misuses of personal information that threaten today's unwitting subjects of social media — and perhaps now scholarly — experimentation and manipulation?

In 2018 this discussion is no longer confined to small groups of research participants. It touches virtually every person who has communicated on social media or a smartphone, taken an online survey, liked a post, or clicked on an ad. Not just the Social Science Research Council and Facebook but everyone should be asking: If we are all research subjects now, what kind of practices and policies will best preserve the values of individual dignity, privacy, and consent?

Given the unique nature of the new collaboration, these questions should be directed to the social scientists who will be making use of novel data sets. But they must also be answered by the corporations and data miners they collaborate with. If the byproduct is a new standard of data ethics with a broad purchase — viewed as the responsibility not simply of academics but also of the multifarious parties now engaged in social and behavioral research — that will truly fulfill the SSRC's mission to "produce findings that improve everybody's lives."

Social researchers rightly seek broad latitude to study social problems. The issues that the Social Data Initiative aims to tackle are, unquestionably, among the most pressing of our time. One imagines that, like Laud Humphreys, its architects believe that the knowledge to be gained is worth the risks. But the cautionary tale of *Tearoom Trade* reminds us that the political terrain can shift beneath researchers' feet. They are not the only arbiters of

what the public, or their own research subjects, will accept. A bold research agenda, even a celebrated one, can swiftly be derailed by ethical missteps. The SSRC-Facebook collaboration might draw that lesson from the 1960s, too.

*Sarah E. Igo, an associate professor of history and director of American studies at Vanderbilt University, is the author of *The Known Citizen: A History of Privacy in Modern America* (Harvard University Press, 2018).*

A version of this article appeared in the October 12, 2018 issue.

Copyright © 2018 The Chronicle of Higher Education