

Protecting human research participants in the age of big data

Susan T. Fiske^{a,1} and Robert M. Hauser^b

^aDepartment of Psychology and Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ 08544; and ^bNational Research Council and Department of Sociology, University of Wisconsin–Madison, Madison, WI 53706

Facebook's experimental manipulation of newsfeed content and the subsequent PNAS publication of significant findings from it (1) have drawn attention to the regulation of human participation in academic research and to the differences between commercial and academic research. Those events were recognized in an Expression of Concern in PNAS (2). In commerce and on the Internet, experimentation is ubiquitous and invisible, and there are no protections for human participants beyond typically unread use agreements. In contrast, academic research is almost always governed by the provisions of the "Common Rule," the US Department of Health and Human Services' Code of Federal Regulations Title 45 Part 46 (45CFR46), "common" because it has been adopted by numerous federal agencies and applied to many research institutions. One might well wonder why academic research is more subject to ethical review than that of business enterprises. Unregulated (Facebook) and regulated (Cornell University) activities were combined in the PNAS publication (1), the former by experimenting with large numbers of unwitting participants, the latter by approving the use of preexisting experimental data as exempt from the university's ethical review.

Time to Revise the Common Rule

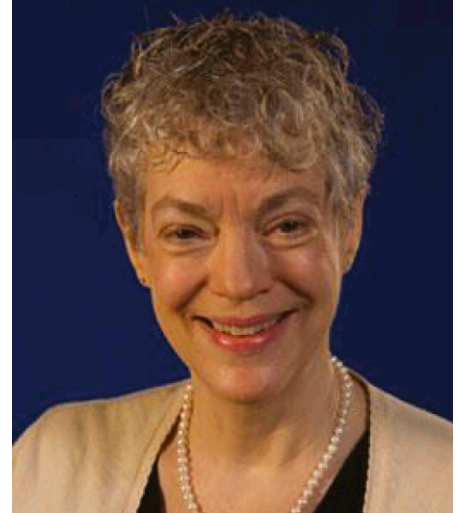
Over the quarter century since the last revision of the Common Rule, the technologies of communication, data collection and analysis, and experimentation have transformed radically. Thus, it is time for a forward-looking revision of the Common Rule that will maintain adherence to the principles of the Belmont report of 1978: respect for persons, beneficence, and justice (www.hhs.gov/ohrp/humansubjects/guidance/belmont.html). Revision began in 2011 with a draft of proposed changes that elicited about 1,000 written comments. Thus, with support from the National Science Foundation and a number of private organizations and foundations, the

National Research Council (NRC) prepared a consensus report on revision of the Common Rule (3).

Among other tasks, the NRC report recommends human subjects regulations for the age of big data. First, the report defines "human subjects research" (HSR) as "a systematic investigation designed to develop or contribute to generalizable knowledge by obtaining data about a living individual directly through interaction or intervention, or by obtaining identifiable private information about an individual" (Rec 2.1) (3). In practical terms, using publicly available information is not HSR, even if information is identifiable, as long as individuals have no reasonable expectation of privacy (Rec 2.3). Examples include observing, coding, and recording behavior in public places (including certain Internet and other digital data) where an individual has "no reasonable expectation of privacy" (3). For example, analyses of posts to a public forum would not be HSR.

The NRC panel recommended adopting the draft regulation's new category of "excused" research for no-greater-than-minimal information risk (Rec 2.5–2.7) (3). That is, researchers could register such projects with their institutional review board (IRB). The relevant IRB would have a short, defined period to object, and without an objection the research could proceed. "Excused" projects would have to fit the standard of no-greater-than-minimal risk (i.e., everyday risk), of which IRBs would audit a small random subset. The main issue would be identity protection, established by registering a privacy-protection plan with the IRB. These recommendations would excuse the reuse of much preexisting data, even with private information, as long as participants' identities are protected.

The NRC panel recommended extending the "excused" category to include benign interventions or interactions that are familiar in everyday life (educational tests, surveys, focus groups), even if the research queries people's



Susan T. Fiske, chair of the National Research Council Committee on Revisions to the Common Rule for the Protection of Human Subjects in Research in the Behavioral and Social Sciences, and Eugene Higgins Professor of Psychology and Professor of Public Affairs at Princeton University.



Robert M. Hauser, Executive Director of the Division of Behavioral and Social Sciences and Education at the National Research Council, and Vilas Research Professor and Samuel Stouffer Professor of Sociology, Emeritus, at the University of Wisconsin–Madison.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. Email: sfiske@princeton.edu.

physical or psychological well-being, as long as participants agree to participate and their identities are protected. Certain public Internet interactions would be included here if they fit the rest of the guidelines.

Understanding Risks in Daily Life

Ambiguity remains, but research can resolve it. To protect human subjects more effectively, according to the no-greater-than-minimal-risk standard, the panel recommends first understanding the risks in daily life of the general population. Otherwise IRBs, researchers, and the public operate on anecdote and hunch. Second, IRBs and researchers need standards “for calculating risk from both the probability and magnitude of harm” (3). HSR protection need not overreact to vanishingly small probabilities of worst-case scenarios (nor underreact to highly probable, greater-than-everyday risk). Third, the panel suggested research on “minimizing potential harms to no-more-than-minimal risk.” Finally, research should “study effects of social and behavioral research on research participants for evidence-based assessments of ‘known and foreseeable’ risk” (3). Research would help take public reactions to the Facebook study out of the realm of speculation and into the realm of evidence.

Given the rapid change in information and technology, ongoing research needs to study

(i) innovations in the data use of nonresearch information and records, (ii) new ways of collecting and linking data, and (iii) new methods for measuring informational risk and risk reduction. Ultimately, research needs to test disclosure-limitation mechanisms against actual datasets to develop best practices and to develop disclosure risk-assessment and risk-mitigation strategies, consistent with “big data” used in the social and behavioral sciences.

Future academic studies (covered by the Common Rule) should carefully tailor consent processes to the relevant context and population, not just use standardized, all-purpose forms (Rec 4.1, 4.2) (3). The consent process should not be used to limit institutional or sponsor liability (Rec 4.3) (3).

A Multifaceted Approach

IRB review does not apply to Facebook and other private enterprises, yet they generate data that can benefit humanity. Reuse of those data (or any Common

Rule-covered data) requires an array of data-protection approaches (Rec 5.1), such as: (i) a *portfolio approach* considering safe people, safe projects, safe data, safe settings, and safe outputs; (ii) a range of statistical methods to reduce disclosure risk; (iii) consulting resources and data protection models, such as university research data-management service groups, individual IT/protection experts, and specialized institutions; (iv) use of existing standards for data protection promulgated by the National Institute of Standards and Technology; and (v) developing a national center to define and certify information risk of different types of studies and corresponding data-protection plans to minimize risks (Rec 5.2) (3).

Human subjects protection is an enduring value. It is especially important that the Belmont report principle of respect for persons—autonomy and protection—must prevail in the age of big data.

1 Kramer AID, 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 Verma IM (2014) Editorial expression of concern: Experimental evidence of massive scale emotional contagion through social networks. *Proc Natl Acad Sci USA* 111(29):10779.

3 Committee on Revisions to the Common Rule for the Protection of Human Subjects in Research in the Behavioral and

Social Sciences; Board on Behavioral, Cognitive, and Sensory Sciences; Committee on National Statistics; Committee on Population; Division of Behavioral and Social Sciences and Education; 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).