

# Brain Schrag's Commentary on "Conducting Research in Online Communities"

Commentary On  
Conducting Research in Online Communities

This case raises issues in research ethics which are in part old and in part new. On the one hand there is the old issue of whether it is ethically justified to do observational research on "public" human behavior. A newer question is whether listserves on the internet are "public" spaces and whether there are privacy norms that are applicable and place restrictions even in "public" spaces. I will address three issues in this case.

## I. Is this Human Subjects Research?

The first issue is whether or not the research proposed in this case is properly classified as human subjects research as defined in the Code of Federal Regulations and therefore falls under the United States guidelines for human subjects research.

In Part 1 of this case, Dr. McIntosh's first suggestion is that Roger simply lurk on line as an unregistered guest and do his research, since the web site can be accessed by unregistered guests to read current postings and archived postings. Is this human subjects research? The relevant guidelines are found in the Code of Federal Regulations PART 46, PROTECTION OF HUMAN SUBJECTS. (CFR 46.102 (1), (2) provide the definition and (CFR 46.101 (b) (2) (4) identifies the activities that are exempt from human research guidelines. **(Code of Federal Regulation)**

Consider first the definitions of human subjects research in the Code of Federal Regulation. (PART 46, PROTECTION OF HUMAN SUBJECTS §46.102 Definitions):

(f), Human *subject* means 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.

*Intervention* includes both physical procedures by which data are gathered (for example, venipuncture) and manipulations of the subject or the subject's environment that are performed for research purposes. *Interaction* includes communication or interpersonal contact between investigator and subject. *Private information* includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example, a medical record). Private information must be individually identifiable (i.e., the identity of the subject is or may readily be ascertained by the investigator or associated with the information) in order for obtaining the information to constitute research involving human subjects.

In Part 1, if Roger merely lurks on line, observing postings or looking at archived postings, there is presumably no *interaction* between Roger and the members of the group; neither is there any *intervention* since he is not manipulating the subject or the subjects' environment. The information Roger would obtain is available to anyone who accesses the web site as an unregistered guest. It is reasonable to say the information is as public as that in a daily newspaper. Given these considerations, it is reasonable to say that the research proposed in Part 1 does not constitute Human Subjects research under the U. S. Code of Federal Regulations.

In Part 2, an alternate proposal to download only the site's archived messages posted the previous year is considered. The forum moderator indicates that participants did not expect at the time of posting that their messages would be used for research purposes; that most participants were unaware that their postings are publicly available and that they view their messages as private communications to other members of the NFF forum. The moderator will only help Roger if Roger first seeks permission from the entire NFF support group.

The fact that participants are unaware that their postings are publicly available and that they consider them private, does not change the status of the research under the definition. This research still does not fall under the definition of Human Subjects

research. For a systematic discussion of what counts as human subjects research on the internet under the Code of Federal Regulation see **(Walther)**.

In Part 3, Roger decides to post a message to the community to inform them that he would like to conduct research on NFF's activities during the next year. At that point Roger has begun to interact with the group. By making the group aware that it is possible for them to be monitored, Roger has destroyed their illusion that this is a private space and may make them self-conscious about their postings. In that sense, Roger has intervened in the group and perhaps already altered the group behavior, whether or not they give him permission to proceed. If they consent to the research, the same research activity which did not previously fall under the definition of Human Subjects research certainly does now.

#### *Impact of human subjects research on group function*

The primary function of this website is to provide mutual support for a group of persons who are concerned with a disorder which is surely distressing to those who have it or those with a loved one with the disorder and all of whom are aware of the social stigma associated with the disorder. Their focus is understandably on that situation and they may deliberately keep access boundaries minimal in order to encourage those who seek support to join in. Their focus is thus inward on their group, not a wider public of strangers scrutinizing their every word. It is not credible to assume group participants would not be affected by the realization that a complete stranger was observing them and reporting their interactions to a larger world of strangers. In this sense, this particular group differs from other groups such as those involved in Face Book where participants assume that what they write is for a wider public.

Thus one effect Roger and his mentor may have in doing their study of this group as proposed in Part 3 is to undermine the function of the group. The group has been created as a support group. That presupposes its members share a common concern and develop a certain trust and a climate of mutual support among the members. It has to be disruptive to be aware that a stranger, who does not share those concerns, is observing and at some point reporting their interactions. **(Elgesem)** That is so whether their anonymity is protected or not. (Imagine the impact on group dynamics if an Alcoholics Anonymous group was aware that a researcher was sitting in on their meeting and would be reporting their discussions and interactions to a wider audience.)

Undermining this narcolepsy group's function as a support group is a moral harm. How does Roger justify the moral harm done to this group in order to observe them? What is the research value of this study that is so important that it justifies undermining the very purpose of the group's existence, especially since there may be many other groups Roger could study for which this is not a consideration?

*Is this research subject to research guidelines of other countries?*

Although it is understandable to approach this case from the perspective of the U.S. Code of Federal Regulation, it is worth noting that since postings in this case could have been made by citizens in other countries, (and Roger has no way of knowing the background of the group) it is quite possible that participant's perceptions of the ethical acceptability of "research by lurking" and the research guidelines of those countries having to do with invasion of privacy in research on human subjects could be quite different than that reflected in the U.S. Code. It is worth noting that European research guidelines are much more inclined to assume a deontological emphasis on rights of individuals and not (as does the U.S. code) allow utilitarian considerations of benefits to others to override those values **(Ess)**.

*Is this Observation of Internet Behavior, the Observation of Public Behavior?*

Whether the research proposed in either Part 1 or Part 3 of this case falls under the definition of Human subjects research is one issue. However, even if, in Part 3 of the case, the research does fall under the definition of Human Subjects research, it may be exempt from humans subjects research guidelines because, it could be argued, it involves observation of "public" behavior.

The exemptions from human subjects research Guidelines are specified in the Code of Federal Regulations §46.101 (b):

§46.101 (Code of Federal Regulation)

(b)...research activities in which the only involvement of research subjects will be in one or more of the following categories are exempt from this policy:

(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:

(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and  
(ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

(4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

Assuming Roger can protect the identity of persons in the group, then if the behavior being observed is in public or the archives of group discussions are public, then Roger may be exempt from human subjects guide lines and is not required by the regulations to obtain informed consent from the subjects. (I hasten to add that, even if the research is exempt, there may still be good reasons to bring this research before an IRB for their review.)

It is an open question whether, in fact, Roger can protect the identity of the participants in doing this research. Bruckman details just how difficult it is to disguise and protect the identity of subjects in research in this sort of online setting.  
**(Bruckman)**

## **II. Is the Internet in the Public Sphere?**

This case raises issues beyond that of simply asking what the Code of Federal Regulations would require of researchers. It raises a new conceptual issue that has implications for this sort of research; that is, should we view the narcolepsy listserve as a public space or public sphere? For the purposes of research ethics, is this listserve a public space, or is it relevantly analogous to a public space, or is it something altogether different? In doing research ethics on the web, it is common to assert an analogy between a public space and space in a public chat room and therefore between observation of subjects in a public sphere and observation of online behavior. **(Ess)** The issue of whether there is an exemption in this research to human subjects guidelines as discussed above does assume that such spaces on the internet constitute a public space. Is that really so?

If the listserv literally is a public space, then there is a case to be made that ethical guidelines regarding observing human action in that “internet space” is no different than guidelines for observing human behavior in the public square. The behavior is thus public and in that sense “up for grabs.” Anyone is free to observe anyone else in the public square and since a listserv is just a variant on the public square, no listserv participant can complain about being the object of surveillance in that setting.

The paradigm of a public space is a public square, with actual people walking about, observable to all, perhaps with people sitting at open air cafes holding conversations accessible to others at nearby tables. Yet, on the internet, people are not in the same physical location, not in visual contact, perhaps not even in the same temporal coordinates. What is true is that the internet is a “technically accessible medium.” But why should technical accessibility be equated with being in the public space? **(Berry, 2004)**

Walther, for example, presupposes that a listserv is literally a public space or relevantly analogous to a public space in which participants cannot reasonably expect that what they say and do should be treated as private. As he notes, research use of conversation, if gathered in a publicly accessible venue is not human subjects research by definition and is parallel to recording conversations in a public park. Collection of data which is publicly available is analogous to collecting data from old newspapers or public broadcasts. **(Walther, p. 207)**

Bruckman, on the other hand, challenges that analogy and argues that our intuitive notions of “public” and “private” in this context can be misleading, and that a web page is neither a public place, like an art gallery nor a private place like one’s home—it is a web page. Bruckman argues that, in thinking about research on the postings on the narcolepsy listserv, rather than invoking the analogy of a person in a public square or public park, perhaps the appropriate analogy is an author of a published work. **(Bruckman)** Is it conceptually clearer to think of the internet as a “space” in which embodied persons “interact,” or is it conceptually more appropriate to think of the internet as a textual repository where authors deposit their work? **(Berry)**

It is true that we do argue that a letter to the editor of a newspaper, addressed to fellow citizens is a public document in the “public” sphere. But, in that setting, the letter is intentionally addressed to a wide audience of strangers. How is that

comparable to what is written by members of the narcolepsy group? What is distinctive of their writings is that they are written explicitly to their group members who share a fairly narrow set of therapeutic goals; they are not writing to the universe as a whole.

If one thinks of postings on the internet as creative writings of authors, then that shifts the use of those materials from a focus on human subjects research guidelines to the permission of the use of copyrighted material. One effect of that shift in paradigm is to force a recognition that such postings, even if public, are not simply “up for grabs” as taped conversations in the park by the researcher may be, but must be treated as copyrighted material.

It is beyond the scope of this commentary to resolve the issue but perhaps it is enough to raise the issue to caution researchers not to simply assume that the internet is a public space and all the usual understandings regarding doing research in public spaces apply.

### **III. Is there no privacy in the public sphere?**

There is a larger issue that goes beyond the question of whether this research is activity in the public sphere. Suppose we grant for the moment that the activity of the narcolepsy group falls within the category of the public sphere. There is a yet more fundamental question to address. That is the question of whether there can be privacy in the public sphere. That possibility challenges the very presupposition of the conventional public/private distinction.

The conventional wisdom, which underlies longstanding practice in observational research in the social sciences and the Code of Federal Regulation human subjects research guidelines, is as Helen Nissenbaum puts it:

If you have chosen to expose yourself and information about yourself in public view with the result that others have access to you or to information about you without intruding upon your private realm, then any restrictions on what they may observe, record or do with this information cannot be justified. **(Nissenbaum, 1998 p. 572)**

This is not an issue unique to the internet. In the social sciences, there is a long history of assuming that public behavior is fair game for observational research and

that there is implicit consent in a subject's public behavior that such behavior may be studied by others. Lurking on the internet, in this case, may be no different than anthropologists observing and writing about the behavior of an isolated, indigenous tribe without the tribe's knowledge or consent or the infamous case of the observational research in the Tearoom Trade case. In all such cases, subjects may be unaware that their public behavior is being recorded and reported to a wider audience of complete strangers.

Nissenbaum and other scholars (**Nissenbaum, 97, 98, 2004; Rachels, 1975; Scanlon, 2001; Schoeman, 1984**) have begun to challenge this conventional wisdom and argue for a fundamental rethinking of the public/private distinction and argue for the notion of a sphere of privacy in public. Nissenbaum has been at the forefront of that discussion as it relates to the internet.

We cannot rehearse the entire argument for this perspective but the basic argument is this. We all live our lives in multiple contexts, realms or spheres, including such contexts as our work setting, visiting friends, seeking health care, shopping, banking and walking the public streets. Each of these contexts is governed by norms, including norms for the exchange of information. The central point is that there is no place that is not governed by informational norms. The notion that when one ventures out in public no norms are in operation is simply pure fiction.

Nissenbaum posits two forms of informational norms for these contexts. One is a norm of appropriateness. This norm dictates that information which may be appropriate and fitting to reveal in one particular context may not be fitting and appropriate to reveal in another context. The kind of information appropriately shared by a patient with a doctor is not necessarily the kind of information that would be appropriate for a doctor to share about himself with a patient. Information on one's financial standing may be appropriately shared with a bank but not necessarily appropriately shared by the banker with acquaintances. It is understood to be inappropriate to take information that is appropriate in one context, e.g. revealing information about oneself in a group therapy session and insert it into another context-e.g. a researcher sharing that information in a research project.

The second is a norm for distribution or transfer of information. We recognize that there are norms regarding the flow of information about ourselves. It is expected that if one shares information with a friend, it would be a violation of the norms of friendship for the friend to share that information with strangers. It would be a



violation of the norms of support groups if information revealed about oneself in that context were to be transformed by someone else into data for their research paper.

On this view then, as Nissenbaum put it,

personal information revealed in a particular context is always “tagged” with that context and never “up for grabs” as other accounts would have us believe of public information gathered in public places. **(Nissenbaum, 2004, p. 121)**

It is beyond the scope of this commentary to assess Nissenbaum’s analysis but if her analysis is right, it does help one to see why there may be a difference between information about a person being “technically accessible” on the internet and a researcher being morally justified in appropriating that information. The mere fact that such information is “public” in the sense of “technically accessible” does not justify its acquisition and use by a researcher. That is what is wrong with Roger lurking on line and using the data from the narcolepsy group for his research. That is what is wrong with downloading the archived data without their consent and perhaps what is wrong with him even approaching them for consent. Nissenbaum’s analysis also raises questions about the general practice, particularly in the social sciences, of research involving observation of human behavior in a public setting. None of this is captured by the current Code of Federal Regulations research guidelines and may call into question the adequacy of those guidelines.

## References

- Bassett, Elizabeth H. and O’Riordan, Kate. “Ethics of Internet research: Contesting the human subjects research model” *Ethics and Information Technology*, Volume 4, Number 3 2002, pp. 233-247.
- Berry, David M. “Internet Research: Privacy, Ethics and Alienation: an Open Source Approach,” *Internet Research*, Vol 14, No. 4, 2004, pp.323-332.
- Bruckman, Amy. “Studying the Amateur Artist: A perspective on Disguising Data Collected in Human Subjects Research on the Internet,” *Ethics and Information Technology*, Volume 4, Number 3, 2002, pp. 217-231.
- Capurro Rafael and Pingel, Christoph. “Ethical issues of Online Communication Research,” *Ethics and Information Technology*, Volume 4, Number 3, 2002, pp.189-194.

- Code of Federal Regulations, TITLE 45, PUBLIC WELFARE, DEPARTMENT OF HEALTH AND HUMAN SERVICES, PART 46, PROTECTION OF HUMAN SUBJECTS <http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm#46.101>. November 17, 2005.
- Ess, Charles. "Introduction" *Ethics and Information Technology*, Volume 4, Number 3 2002, pp. 177-188.
- Elgesem, Dag. "What is special about the ethical issues in online research?" *Ethics and Information Technology*, Volume 4, Number 3 2002 pp. 195-203.
- Nissenbaum, Helen. "Toward an Approach to Privacy in Public: Challenges of Information Technology," *Ethics and Behavior*, 7(3) 1997, pp. 207-219.
- Nissenbaum, Helen. "Protecting Privacy in an Information Age: The Problem of Privacy in Public," *Law and Philosophy*, 17: 1998, pp. 559-596.
- Nissenbaum, Helen. "Privacy as Contextual Integrity," *Washington Law Review*, Vol. 79, No. 1, 2004, pp. 101-139.
- Rachels, James. "Why Privacy is Important," *Philosophy and Public Affairs*, 1975, 4: 323-33.
- Scanlon, Michael. "Informational Privacy and Moral Values," *Information Technology*, Volume 3, 2001, pp. 3-12.
- Schoeman, F., (ed.). *Philosophical Dimensions of Privacy: An Anthology*, Cambridge: Cambridge University Press, 1984.
- Walther, Joseph B. "Research ethics in Internet-enabled research: Human subjects issues and methodological myopia," *Ethics and Information Technology*, Volume 4, Number 3, 2002, pp. 205-216.