Doing Research in Cyberspace

David Jacobson, Field Methods (Vol. 11, No. 2, November 1999)
pp. 127-145, copyright ©1999 by Sage Publications, Inc.
Reprinted by Permission of Sage Publications, Inc.

Doing Research in Cyberspace

DAVID JACOBSON
Brandeis University

Although laws and policies developed in the context of off-line research apply to online investigations, questions about the identifiability of human subjects, the conceptualization of privacy, the need for and means of obtaining informed consent, and the applicability of copyright law to computer-mediated communication (CMC) pose special problems for doing research in cyberspace.

In the past several years, there has been a spectacular expansion in use of the Internet and its various modes of computer-mediated communication (CMC), commonly referred to as cyberspace. It is difficult to get precise figures on the size of the Internet, but the estimates are impressive. One source reports almost 30 million computers (i.e., hosts connected to the Internet in January 1998, an increase of almost 50% from the preceding year and a 22-fold increase from January 1993; Zakon 1998). The World Wide Web (WWW) has also experienced explosive growth: In April 1998, there were 2,215,195 sites, almost double the number from the year before (Zakon 1998).[1] Although the number of people who connect to or use the Internet is undetermined, it is estimated that in February 1998, there were 62 million users in the United States, and in July 1998, there were 130 million users worldwide (see NUA Surveys 1998).

In keeping with these phenomena, there has been a dramatic increase in research about the ways in which users of cyberspace act and interact. This work has been conducted by social scientists, scholars in the humanities, and legal scholars, among others (e.g., Spears and Lea 1994; Lea and Spears 1995; Marvin 1995; Turkle 1995; Herring 1996; Jacobson 1996; Mnookin 1996; Parks and Floyd 1996; Walther 1996; Parks and Roberts 1997). At the same time, social scientists and others have begun to express concern about

I wish to express appreciation to several people for discussions we had during the time I was researching the issues addressed in this article and for their comments on an earlier draft of it. My thanks to Amy Bruckman, Robert Canary, Mary Davis, Samantha Herdman, Elizabeth Hess, Lois Jacobson, Roberta Nary, Mac Parks, Zick Rubin, Judith Sizer, Charley Stivale, Kaisa Vahahyyppa, and Joe Walther. I would also like to thank Russ Bernard and the anonymous reviewers from Field Methods; their comments helped strengthen the argument presented here.

Field Methods, Vol. 11, No. 2, November 1999 127-145

several issues in online research (see Branscomb 1996; Thomas 1996; Bruckman 1997). Most of this discourse has focused on the protection of human subjects and related matters. Although laws and policies developed in the context of off-line research apply to online investigations, questions about the identifiability of human subjects, the conceptualization of privacy, the need for and means of obtaining informed consent, and the applicability of copyright law to CMC pose special problems for those doing research in cyberspace. My goal in this article is to consider issues that bear on the development of guidelines for doing this sort of research and to suggest ways of managing the problems involved in conducting studies of online behavior.

However, a caveat is in order. The Internet spans national boundaries, and researchers (and those whom they research) come from many nations other than the United States. Their work may not fall within the jurisdiction of U.S. rules and regulations. The argument of this article dealing with issues of human subjects pertains to research “supported by or subject to regulation by any Federal Department or Agency” (see Code of Federal Regulations [CFR] 1991, Sec. 46.101 [a]). Researchers employed by organizations, including universities, supported by U.S. federal funds are subject to these regulations. Similarly, the discussion of privacy and copyright issues is also informed by (and limited to) U.S. law. It is beyond the scope of this article to consider matters of international and foreign law.

Before examining the applicability of off-line regulations to research in cyberspace, a brief note about various modes of CMC is in order. CMC may be classified as asynchronous or synchronous. Asynchronous CMC includes e-mail, mailing lists, bulletin board services and newsgroups, and Web resources (including pages and sites; see Gray 1996) where messages are not exchanged in real time.’ In such asynchronous systems, the messages are often permanent (unless deleted) and may be stored in some form of archive. E-mail differs from other forms of asynchronous CMC in that it is a point-to-point system involving the exchange of messages between individual senders and receivers (although there may be more than one recipient when the message is copied to others or sent to multiple recipients in a distribution file), whereas bulletin boards and newsgroups involve the posting of messages by individuals that are read by anyone who connects to them. (There is a technical difference between newsgroups and bulletin board systems: In the former, multiple copies of each message are distributed widely on different machines; in the latter, messages are posted only to a particular server to which people connect.) Mailing lists combine features of e-mail and newsgroups: They distribute messages, via e-mail, from an individual sender to all who subscribe to a list. In synchronous types of CMC, including chat rooms provided by

Internet service providers and different types of virtual realities (e.g., MUDs, MOOs; Curtis 1992), users communicate in real time.' The messages are transient and, unless recorded, do not endure beyond the time when they are created, sent, and received. However, transiency in synchronous CMC does not preclude fixation, one of the criteria for the establishment of copyright, a point to which I return below.

OFF-LINE RESEARCH GUIDELINES

Much social science research is subject to federal regulations, specifically the CFR (1991) Title 45, Part 46 (Protection of Human Subjects) and the National Institutes of Health's (NIH's) (1993) *Institutional Review Board Guidebook.* For most social scientists, the relevant regulations are those pertinent to research entailing "survey procedures, interview procedures or observation of public behavior" (CFR 1991, Sec. 46.101 [b]) and involving human subjects. *The* CFR *(Sec. 46.102 [f])* defines a *human subject* as "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." According to the regulation (CFR 1991, Sec. 46.102 [f] [2]), 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." Regulations regarding human subjects apply when "information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects" (CFR 199 1, Sec. 46.101 [b] [2] [i]). Exempt from these regulations is

> research involving the collection or study of existing data, documents, records. . . 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. (CFR 1991, Sec. 46.101 [4])

IDENTITY ANONYMITY; AND PSEUDONYMITY

Anonymity and pseudonymity complicate the issue of human subjects in cyberspace. Reviewing the relevant federal guidelines, one researcher (Thomas 1996: 113) concluded "it is indisputable that cyberspace research entails 'human subjects.' " Although that conclusion is true, the situation is more complex. Recall that a human subject is defined as a person whose identity "is or may readily be ascertained by the investigator or associated with the information" (CFR 1991, Sec. 46.102 [f]) about the individual. Although identity is not defined in the relevant federal regulations, it would seem to refer to an individual's actual identity, physical identity, or real identity (see Froomkin 1995, art. 4, par. 1 l), not to a person's social identity or, in the- context of cyberspace, digital persona (Froomkin 1995, par. 33; see Tien 1996). In this sense, research involving individuals whose true identities are not known or cannot be readily ascertained would be exempt from the federal regulations governing human subjects.

Identity in cyberspace is, in a significant sense, indicated by an e-mail address, the header of which normally includes an individual's log-in name and Internet provider (IP) address (see Donath 1996; Gray 1996; Sewell 1996). However, it is technically possible to remove such information from e-mail (and from messages posted via e-mail to mailing lists, bulletin board services, and newsgroups), rendering the individual who sends it unidentifiable. This can be accomplished through anonymous remailers (see Detwiler 1993; Long 1994; Bacard 1995; Branscomb 1995; Froomkin 1995; Lee 1996; Post 1996b, 1996c; Tien 1996). Remailers work by stripping identifying information from the header of an incoming e-mail and forwarding it to its intended recipient with a substituted header that disguises the original sender's name and electronic address.

The use of public key encryption systems, independently and/or in conjunction with anonymous remailers, reinforces anonymity in cyberspace (see Commonwealth of Massachusetts 1996; Froomkin 1995; Information Infrastructure Task Force 1995; Rosoff 1995; Post 1996a, 1996c; Rosenoer 1997). A public key encryption system, such as Pretty Good Privacy (PGP), is a software program that uses an algorithm to generate two keys (or codes) to encrypt (scramble) and decrypt (unscramble) messages. One is a public key, widely publicized by its owner and used to encrypt messages sent to him or her. The other is a private key, kept secret by its owner and used to decrypt them. A message encrypted with a public key cannot be unscrambled without the private key.

Just as anonymity conceals an individual's real identity in online communication, pseudonymity also serves to disguise it. The techniques that make it possible to send messages anonymously can also be used to create pseudonymous messages. Thus, an individual can use an anonymous remailer to send e-mail and/or post messages to bulletin boards and newsgroups using a pseudonym or nom de plume rather than a real name (see Froomkin 1995). The remailer removes the original header, substitutes a pseudonym, and forwards the message to its ultimate destination. Conversely, others may respond to the e-mail or posting by replying through the remailer, which replaces the pseudonymous header with the original sender's real name and address and forwards it to her or him.

The use of pseudonyms is widespread in chat room and virtual communities. In most MOOs, for example, participants typically assign pseudonyms to their online characters (see Curtis 1992; Jacobson 1996). The name, gender, and other characteristics of a participant's digital persona may be completely unrelated to the individual's off-line identity. As Mnookin (1996) remarked about the virtual community on which she worked, "A Lambda-MOO character need in no way correspond to a person's real life identity; people can make and remake themselves, choosing their gender and the details of their online presentation; they need not even present themselves as human." And although system administrators in virtual communities (as do administrators of other modes of CMC) may have access to the off-line identities of participants, or at least to their names and e-mail addresses, ordinary participants (including researchers) do not.[4] Furthermore, where an e-mail address is known, it may be that of an anonymous remailer, as described above, and, in such cases, even system administrators may not have access to a participant's true identity.

Moreover, system administrators are barred by privacy laws from divulging that information, except under certain circumstances. Under the provisions of the Electronic Communications Privacy Act of 1986, it is illegal to intercept any electronic communication in transit or to divulge the contents of stored electronic communications unless (1) the provider of an electronic communication service is engaged in an activity for which it is necessary to provide the service or to protect the integrity of the system, (2) it is required by a court order to assist lawfully authorized persons to intercept electronic communications, or (3) a user of the service waives his or her rights to privacy (see Cavazos and Morin 1994; Branscomb 1995; Rosenoer 1997). In addition, the local policies of many virtual communities reflect the privacy laws. For example, although most MOOs indicate that system administrators have access to everything communicated in these online communities, including the off-line e-mail addresses of their participants, they also note that administrators will be discreet in their management of such information. At least one virtual community goes further in acknowledging the rights and responsibilities of its system administrators regarding the privacy of its participants. An administrator at Diversity University, an educational MOO, posted a message (Message 1738 on *General, a mailing list internal to it, dated February 24, 1997, MOO.dumain.duets.org) that reads in part,

> By using DU MOO, as specified in the splash screen, you indicate your understanding that the administration may monitor communications. DU administrators monitor communications specifically when someone uses the @witness command, indicating that someone is by their misbehavior depriving them of the right to an atmosphere conducive to teaching, learning, and social services. DU administrators never randomly monitor communications, as that would be against the law. In addition, planting of undisclosed listening devices or other illegitimate recording of other people's private conversations is sufficient to lead to the offender losing access privileges.

Despite the barriers posed by anonymity and pseudonymity, researchers are not necessarily exempt from the regulations governing human subjects since it is possible, under certain circumstances, to uncover an individual's real identity. For example, people may (and do) voluntarily disclose identifying information about themselves in the course of developing personal relationships in cyberspace (see Allen 1996; Jacobson 1996). And it is not uncommon for participants in virtual communities to attempt to discover one another's true identities by collating various bits of information dispersed within a single community. As Allen (1996: 184) notes, "I have found that participants actively seek to erode the anonymity of others, piecing together mosaics of information casually revealed over time or sought out using surveillance tools in the site to triangulate on the identity of others."

It is also possible to seek identifying information across virtual communities. For example, some MOOs enable their participants to remain anonymous, using whatever pseudonyms they prefer and concealing their log-in names and sites. Others permit the use of pseudonyms but make log-in information available, including a participant's e-mail address. These data may be accessed by users typing requests for information about characters. Since it is not uncommon for people to have characters on several MOOs and to use the same character name or pseudonym on each of them (thereby contributing to the establishment of a digital persona), a researcher (or any other participant) would only have to track a character name from an anonymous MOO to one that was not in order to discover off-line information about the person using that pseudonym. Moreover, since a person may use the same pseudonym across different modes of CMC (e.g., newsgroups, chat rooms), it may be possible to develop identifying information by comparing these different places or spaces. On the other hand, different participants may use the same pseudonym in different virtual locales, a practice that could confound the identification process.

However, the typist problem complicates efforts to discover the off-line identity of participants in cyberspace, even in situations in which the name and electronic address of a digital persona are readily ascertainable. The typist is the actual human being, the person at the keyboard, who writes the e-mail, posts the message to a newsgroup, and/or controls the character in a virtual community; the problem is that the identity of the typist cannot be directly and unambiguously inferred from the e-mail address or from the text she or he produces. Several circumstances contribute to this uncertainty. One is the unauthorized or illegal use of a computer account, as Detwiler (1993, par. 1.5) states:

> Currently intemet users do not really have any great assurances that the messages in e-mail and USENET are from who they appear to be. A person's mailing address is far from an identification of an individual. Anyone with access to the account [legally or otherwise] can send mail [from it].

Another is pooled computer facilities (see Branscomb 1995) or shared accounts (see Tien 1996), a situation in which two or more potentially unidentified individuals legitimately use the same terminal and/or computer account. For example, in his analysis of a "rape" in the virtual reality of LambdaMOO, Dibbell (1998:30) notes that the account used by the perpetrator "had been the more or less communal property of an entire NYU dorm floor." As Allen (1996) observed, participants in virtual communities recognize the difficulty of establishing a verifiable link between typist and character and understand it offers them plausible deniability for their actions and interactions.

The typist problem is particularly significant for social research in cyberspace. It not only complicates the application of guidelines governing the protection of human subjects, as I have argued, but it also raises a question about online studies of gendered behavior and other related topics when the identities and personal characteristics of participants are not determined independently of the ways in which they are represented in cyberspace. In fact, there are two aspects to this question. The first is logical: We cannot know with any certainty who is at the keyboard and therefore there will always be doubt, if online research is not supplemented by off-line research, about precisely who is sending an e-mail message or occupying a character in a virtual reality.' The second issue is empirical: How widespread is the practice of misrepresentation of gender (and/or other sociodemographic characteristics) in CMC? Although the answer to that question may be decided by a combination of online and off-line inquiry, in fact, most research in cyberspace has not obtained and/or does not present data on the off-line (or "real") characteristics of its participants (respondents or informants), and the conclusions they reach (e.g., in cyberspace, women behave differently in various ways than do men) are unwarranted by the evidence provided.

PRIVACY

To the extent that individuals in cyberspace can be identified, researchers must be concerned with matters of privacy. In the CFR (1991, Sec. 46.102 [f] [2]), the principle of a reasonable expectation of privacy is evident in the definition of private information, which "includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place." One issue that concerns both researchers and participants in cyberspace is the conceptualization of such contexts (see Allen 1996; Gajjala 1996; King 1996; Waskul and Douglass 1996; Bruckman 1997). According to the *Institutional Review Board Guidebook* (NIH 1993, chap. 3, p. 30), "pedestrians on the street" would be an example of a public place and "people in their homes" would be an example of a private place, there being a reasonable expectation of privacy in the latter but not in the former. However, it also holds that "Some behavior that occurs in public places may not really be public behavior-the individuals involved have a reasonable expectation of privacy." The guidebook cites an example of a conversation in a public park, presumably between a couple or among a few friends sitting on a bench (see Waskul and Douglass 1996).

This type of situation is especially relevant to research in cyberspace. Chat rooms and virtual communities are public places in the sense that anyone with access to the Internet may connect to them, either as a temporary visitor or as aregistered member-and in most instances, there are no restrictions on membership. However, a distinction is made between public and private spaces, the former being rooms anyone may visit anytime and the latter being rooms created by individual participants to which they may restrict access.

Like off-line homes, private rooms are places where participants have a reasonable expectation of privacy. There is, however, no equivalent to a tête-à-tête in a park because all communicative actions in public rooms (other than private messages such as whispers, pages, and remote emotes that are restricted to specified persons) are distributed for all to hear, or, more precisely, read.[6] There is no reasonable expectation of privacy in these conceptual spaces, and federal regulations governing research involving human subjects would not apply to information about behavior in them (unless, of course, participants can be identified).

Another issue concerns the content of CMC. Federal regulations about privacy and private information pertain to "methods used to obtain information about subjects" rather than to the substance of the information itself (NIH 1993, chap. 3, p. 27). This distinction is often overlooked by participants and researchers in cyberspace and is a source of considerable confusion for them (see Gajjala 1996, Waskul and Douglass 1996). Some participants consider messages sent to newsgroups, bulletin board systems, and mailing lists to be private because the communications refer to personal information-for example, statements made in online support groups by people who are victims of domestic violence or sexual abuse. However, contrary to such perceptions, messages posted to publicly accessible fora are not private and are not protected by privacy laws (see Cavazos and Morin 1994; Rosenoer 1997).

CONSENT

When it is determined that research involves human subjects, federal regulations require researchers to obtain informed consent. The basic elements of informed consent are that researchers provide information to (potential) subjects about the aims and methods of the research, the subjects' rights and responsibilities, and those responsible for conducting the research (see CFR 1991, Sec. 46.116). The CFR (Sec. 46.117 [a]) also stipulates that "informed consent shall be documented by the use of a written consent form approved by the IRB [Institutional Review Board] and signed by the subject or the subject's legally authorized representative."

However, there are exceptions to this rule. The regulations permit a waiver of the requirement to obtain informed consent when the signed consent form is "the only record linking the subject and the research and the principal risk would be potential harm resulting from a breach of confidentiality" (CFR 1991, Sec. 46.117 [c] [l]), an exception relevant to research in cyberspace (and other situations) in which subjects wish to maintain their anonymity (see Myers 1987). A waiver of the requirement is also permitted when an IRB determines that "the research presents no more than minimal risk of harm to subjects and involves no procedures for which written consent is normally required outside of the research context" (CFR 1991, Sec. 46.117 [c] [2]). This latter provision is particularly relevant to ethnographic research entailing observation of daily life in public places. Indeed, the *Institutional Review Board Guidebook* recognizes the special circumstances of fieldwork and their bearing on the requirement to obtain informed consent in ethnographic research. It notes (NIH 1993, chap. 5, p. 5) that the processes of fieldwork "involve complex, continuing interactions between researcher and hosts that cannot be reduced to an informed consent form," a perspective endorsed by the American Anthropological Association (see American Anthropological Association, Code of Ethics, 1997, III A.4). The American Sociological Association holds, in its Code of Ethics (1997, Sec. 12.01), that "sociologists may conduct research in public places or use publicly available information about individuals (e.g., naturalistic observations in public places, analysis of public records, or archival research) without obtaining consent." The American Psychological Association (APA) (1998) has a similar stipulation: Research that "does not require informed consent of research participants" includes "anonymous questionnaires, naturalistic observations, or certain kinds of archival research."

Thus, the ethical codes of conduct of social science professional associations, like federal guidelines, do not insist on obtaining informed consent under the kinds of circumstances that characterize much, if not most, research in cyberspace. For example, the members of ProjectH Research Group, an interdisciplinary collection of CMC researchers, adopted the following position: The "issue of informed consent . . . does not apply to the ProjectH quantitative content analysis, as we intend to analyze only publicly available text. We believe posts are public and their use is governed by professional and academic guidelines" (Rafaeli et al. 1998:268).

Nevertheless, when applicable, the requirement to obtain a signed consent form poses a challenge for those doing research in cyberspace. Certainly, a researcher may (and should) go off-line to obtain a participant's signature by conventional means (i.e., on a form signed by the participant or the participant's legally authorized representative) when a participant is willing to provide it (and the additional information implied by or associated with it; e.g., full name, street address, and so forth). However, participants in research projects, even when their e-mail addresses are known to or ascertainable by a researcher, may be unwilling to provide it and/or the identifying information associated with it.

In such cases, researchers will have to resort to other means of computer-mediated documentation. One way is suggested by work on electronic contracts. To be legally enforceable, certain contracts (those falling under the statute of frauds) require the written signature of the parties involved. However, legal experts hold that other means of conveying a written signature may suffice, including a telegram bearing a typewritten name and, in all likelihood, an e-mail (see Cavazos and Morin 1994).

Perhaps e-mail documentation will suffice for the purpose of informed consent, particularly if it has been encrypted in a public key system of the sort described above or uses a digital signature, a procedure that also involves cryptographic techniques (Cavazos and Morin 1994; Froomkin 1995; Information Infrastructure Task Force 1995; Commonwealth of Massachusetts 1996; Ellison 1996; Electronic Privacy Information Center 1998; Rosenoer 1997).[7] One (potential) problem with such encryption-decryption systems is that research participants may not have them or routinely use them.

In addition to the prospect of e-mail documentation (encrypted or not), another possibility entails the concept of implied consent. An example of this would be a research participant who, by the very act of responding to an instrument like a questionnaire, indicates her or his consent. This situation would be analogous to that of a person who, in using the service of an Internet access provider, agrees to the company's terms regarding such use (which are typically specified in the provider's contract), the so-called end-users agreement. A related form of implied consent entails the use of a Web-based form. Potential research participants access a Web page, read an account of the research aims and their involvement in it, and then click on a button on the page that states they have read the information provided and consent to take part in the research.

Still another alternative, especially appropriate to research conducted in chat rooms and virtual communities, involves creating an electronic record. A researcher could electronically record ("log") communication with a potential participant, including information regarding the nature of the research, the expectations and rights of participants, the identity of those responsible for the conduct of the research, and the person's consent to participate in the research. The researcher could make copies of the log, e-mailing one to the participant and keeping one for himself or herself. This method does not have the security of encrypted data-the text can be manipulated, although the duplicate copies may serve as a safeguard.

All of these methods are consistent with the American Anthropological Association's Code of Ethics (1997, III A. 4), which notes that "informed consent, for the purposes of this code, does not necessarily imply or require a particular written or signed form."

COPYRIGHT

Copyright law raises other issues for researchers in cyberspace. Although messages posted to publicly accessible newsgroups and mailing lists and statements made in chat rooms and in public places in virtual reality communities are not governed by human participant regulations or by privacy laws, they are subject to copyright law. Copyright law (U.S. Copyright Office 1996, United States Code, Title 17, Copyrights) protects the property rights of authors or creators in works that are original expressions, including works expressed in words, sounds, and visual images, and are fixed in a tangible medium (see 17 U.S.C. Sec. 101; Cavazos and Morin 1994; Brandriss 1996; Rosenoer 1997). The Copyright Act states that "a work is fixed in a tangible medium when its embodiment in a copy or phonorecord, by or under the authority of the author, is sufficiently permanent or stable to permit it to be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device, for a period of more than transitory duration" (17 U.S.C. Sec. 101; see Cavazos and Morin 1994; Information Infrastructure Task Force 1995; Brandriss 1996). As noted by the Copyright Office (1996),

> Copyright protection subsists from the time the work is created in fixed form; that is, it is an incident of the process of authorship. The copyright in the work of authorship immediately becomes the property of the author who created it.

Before examining the application of copyright law to various modes of CMC, the following scenario of off-line behavior illustrates the concepts of original expression and fixation. Suppose one person writes a letter to another. If the author creates the document, using his or her own words and not copying those of another (and without regard to his or her creativity or to the artistic merit of the letter), it is an original expression. And as soon as the ink dries on the paper, the words are fixed (see Information Infrastructure Task Force 1995). The recipient of the letter owns the paper, but the sender owns the words. The former cannot use the words-that is, quote them-without written permission of the latter, except under the special circumstances of fair use (which will be addressed below).

CMC is protected by copyright. Asynchronous CMC is clearly covered by the law. As Cavazos and Morin (1994:56-57; see Rosenoer 1997) note, "authors of e-mail automatically hold a copyright in their words. The moment your e-mail is fixed on a storage device somewhere, you own it, even if you have not attached a copyright notice declaring your ownership." Similarly, messages sent by e-mail to newsgroups, bulletin boards, or mailing lists are copyright protected (see Cavazos and Morin 1994; Rosenoer 1997).

Synchronous CMC is also protected. When people speak to one another off-line, their words are not fixed unless tape recorded or transcribed (see Cavazos and Morin 1994). By contrast, "talk" in chat rooms and in virtual communities is fixed in the process of being produced. Although the text of the words users type is not ordinarily stored in the database of the virtual community, it exists in the server's random access memory (RAM). The text is also captured in the RAM of users' computers, and communications software packages enable users to record text appearing on screen, either by initiating a logging process before the text is produced or by capturing it from a scroll-back buffer afterward. And courts have found that a work in RAM is sufficiently permanent to permit it to be perceived, reproduced, or otherwise communicated and is therefore fixed and protected by copyright (see Information Infrastructure Task Force 1995:28 [n. 67]; Brandriss 1996:246-51; Rosenoer 1997:2 [n. 7],5 [n. 58,59], 22-26; in these sources, the cases cited *are MAI Sys. Corp. v. Peak Computer* and *Advanced Computer Servs. v. MAI Sys. Corp).*

Although copyright law provides protection for works of original authorship, there are limitations to it that are significant for researchers. The doctrine of fair use permits use of copyrighted material, without the consent of its owner, "for purposes such as criticism, comment, news reporting, teaching, scholarship, or research," and the law stipulates factors to be considered when determining fair use, including "the purpose of the use, the nature of the copyrighted work, the amount and substantiality of the portion used in relations to the copyrighted work as a whole, and the effect of the use upon the potential market for or value of the copyrighted work" (U.S. Copyright Office 1996, U.S.C. 17, Sec. 107; see Cavazos and Morin 1994:5&55; Information Infrastructure Task Force 1995:73-82; Rosenoer 1997:16-19).

Thus, it would seem that researchers would be free from the restrictions of copyright law. However, it should be noted that in litigation concerning fair use, a court has to consider all four factors noted above and, in any given case, that could mean that a particular use of copyrighted material, although done in the context of research, might not be considered to be fair use. In this regard, use of material collected in cyberspace is constrained in ways similar to the use of off-line materials: Just as there is an obligation to obtain permission to include text from another source published in a monograph or a textbook, so there is a requirement to obtain permission to use material available online.

CONCLUSION

The nature of CMC raises issues that are of concern to those doing research in cyberspace. At present, as this article indicates, these issues include the identifiability of human subjects, the conceptualization of privacy, difficulties associated with obtaining informed consent, and the applicability of copyright law. To the extent that the off-line identities of participants in cyberspace *are* not known or ascertainable to researchers, it would appear that federal guidelines regarding human subjects would not apply to such research. Similarly, to the extent that CMC is public, the privacy provisions of these federal guidelines do not pertain to it. On the other hand, to the extent that identifiable human subjects are involved in online research, the issue of informed consent is relevant. It is clear that researchers seeking approval from their institutional review boards will have to comply with the relevant federal regulations, yet it is unclear which particular protocols institutional review boards will accept for obtaining consent in online research. In any event, the issue will have to be decided and there is likely to be considerable discussion in review boards (and/or human subject committees) until old standards are revised or new regulations are established.

However these problems are managed, with technological and legal innovations, the issues facing researchers will change. Developments that enhance security, privacy, and confidentiality in CMC will likely shape not only the behavior of participants in cyberspace but also the possibilities of obtaining informed consent from them. Furthermore, modifications in copyright law are being considered and debated *(see Electronic Frontiers Florida News* 1998; Hardy 1998; U.S. Copyright Office 1998).

The report of the Information Infrastructure Task Force (1995) contains proposals that would strengthen the legal means available to copyright owners to enforce the rights they hold in their protected works. Legal experts (Brandriss 1996; Elkin-Koren 1996) have criticized this position, arguing that the characteristics of cyberspace necessitate a reconceptualization of the terms of copyright law and a restriction in its application to CMC. Whatever the direction of such changes, the issues they raise will have to be addressed in future research in cyberspace.

Finally, there is a question of the relationship between law and ethics. Many online researchers are guided by, or contend that they and others should be guided by, the ethical codes promulgated by their professional associations, although as Paccagnella (1997) notes, there is little agreement among cyberspace researchers regarding ethical guidelines in general and those pertaining to informed consent in particular. Some researchers seek to obtain it, others do not, and many do not indicate whether they have done so. To the extent that an ethical position might be more inclusive than a legal one, researchers might want to formulate and follow ethical guidelines that go beyond such legal bases. For example, they might choose to treat pseudonymous characters as if they were no different from human subjects whose real identities were known or knowable, assigning pseudonyms to pseudonyms (see Jacobson 1996).* However, that too is a question that will have to be addressed in the future when assessing the issues involved in doing research in cyberspace.

NOTES

1. For adiscussion of terms such as hosts and Web sites, see Gray (1996); for adiscussion of ways in which the Internet and its components are conceptualized, see Quarterman and Carl-Mitchell (1994) and Quarterman (1996).

2. Althoughe-mail is an asynchronous modeofcomputer-mediated communication (CMC), two people simultaneously connected to the Internet (or to any network, local or otherwise) may send e-mail messages back and forth immediately on receiving and reading them, simulating real-time communication. However, if they disconnect, their e-mail messages are stored on their computers (or on a server) and are available to be read and answered at another time. This latter feature contrasts with real-time CMC that occurs, for example, in MOOs, MUDS, and other text-based virtual communities, in Internet relay chat (IRC), in chat rooms, and in systems such as America Online's (AOL's) Instant Messenger, which AOL contrasts with e-mail and describes as a kind of chat system (see http://www.aol.com/aim/). In these modes of communication, persons who are simultaneously connected to a network typically exchange typed messages in real time (although lag or disruptions caused by server and/or network overload and/or by participant unresponsiveness may cause delays in the communication flow). Should participants in these "conversations" disconnect, the messages are not saved (see note 3).

3. Synchronous CMC includes applications such as AOL Instant Messenger (AIM), ICQ, and Excite PAL, which arc basically chat programs that enable people to send messages back and forth in real time, although they also permit users to send e-mail and to participate in other forms of CMC (see AIM's Help tile). Virtual communities such as MUDS and MOOs are based on software programs that permit multiple users to simultaneously access a shared database and communicate and interact in a virtual environment. Communication is text based. Despite metaphors of speech and other physical activity, users act and interact by typing: The objects they create and manipulate and the messages they send and receive appear as words scrolling across a screen. For examples of such virtual communities, see Jacobson (1996) and Marvin (1995).

4. In social and educational MOOs, for example, identifying information is available to system administrators. To become a registered player in one of these virtual communities-one who has a permanent object number, a character name, and the capacity to build (in contrast to "guests" who do not have these attributes)-a person must request a character name and provide an e-mail address. The command for doing so is @request <player-name> for <e-mail-address>. MOO system administrators (e.g., "wizards" and, in some systems, their assistants) have access to that information. Indeed, on many MOOs, such information is archived on a list (Player Creation Log). Moreover, wizards can ascertain the place (Internet provider address) from which a registered player is (or was last) connected: The command for doing so is @net-who <who>. (A list of all of the places from which a character has connected is available by using the command <player>.all_connect_places.) Wizards can also obtain similar information for unregistered, nonpermanent players (guests): The command for doing so is @guests, which returns the players' Internet provider addresses. These kinds of identifying information enable system administrators to exercise social control, since they may block connection to the virtual community by a specific player or by all players (characters and guests) from a specific site. Other kinds of virtual communities, including many chat rooms, also obtain identifying information from their users. For example, Cool Chat (http://www.coolchat.com) requires a valid e-mail address from people who want to participate in its chat rooms (see http://www.coolchat.com/membershipl.html), and the virtual communities of Bianca's Shacks (http://www.bianca.com) require a person to provide a name and credit card number to obtain membership (see http://www.bianca.com/

warez/order.html). Of course, virtual communities operated by Internet service providers like America Online have access to the off-line identities of community participants.

5. The typist problem is not peculiar to doing research in cyberspace,occurring off-line in research entailing the use of mailed questionnaires. For example, Bernard (1994:259, 262), in a discussion of the advantages and disadvantages of various survey methods, notes that in "face-to-face interviews" researchers "know who answers the questions," whereas with a "mailed questionnaire," researchers "can't be sure that the respondent who received it is the person who filled it out." (Presumably, telephone interviews, as well as other forms of inquiry in which participants are anonymous, raise similar issues.) Because the technology of CMC allows participants to bc anonymous and because much research in (and commentary on) cyberspace is about personal identity-and gender identity in particular-the question of who is producing the behavior regarded as data is particularly salient. A possible solution to the typist problem involves the use of digital signatures, although they typically apply to documents and not to the kinds of real-time communication that occurs in various kinds of chat environments. Digital signatures may also be problematic without independent confirmation of the person using them (see note 7).

6. For a discussion of the common conventions of communication in virtual communities, see Jacobson (1996).

7. A digital signature uses cryptographic techniques to encode a digital document in order to authenticate it.

> The digital signature serves as a means for authenticating the work, both as to the identity of the entity that authenticated or "signed" it and as to the contents of the file that encodes the information that constitutes the work. Thus, by using digital signatures one will be able to identify from whom a particular file originated as well as verify that the contents of that file have not been altered from the contents as originally distributed. (Information Infrastructure Task Force 1995: 187)

As in the case of people who share passwords (in other security systems as well as those associated with electronic communication and commerce), there is a possibility that others may have access to a digital signature that does not belong to them, thereby underminingthe authentication process.

8. See the arguments of Froomkin (1995), Post (1996c), and Tien (1996) regarding the status of pseudononymous characters and digital persona.

REFERENCES

Allen, C. 1996. "What's wrong with the golden rule?' Conundrums of conducting ethical research in cyberspace. *Information Society* 12(2): 175-87.

American AnthropologicaI Association. 1997. Code of Ethics of the American Anthropological Association. Final draft, March 1. Available from the World Wide Web at: http://www.ameranth assn.org/ethcode.html.

American Psychological Association. 1998. Ethical principles of psychologists and code of conduct. August 5. Available from the World Wide Web at: http://www.apa.org/ethics/code.html#6.12.

American Sociological Association, 1997. Code of Ethics. July 30, 1998. Available from the World Wide Web at: http://www.asanet.org/ecoderev.htm.

Bacard, A. 1995. Frequently asked questions about anonymous e-mailers. July 30.1998. Available from the World Wide Web at: http://www.eff.org/pub/Privacy/Anonymity/anon_re-mailer.faq.

Bernard, H. R. 1994. Research methods in anthropology: Qualitative and quantitative approaches. 2d ed. Thousand Oaks, CA: Sage.

Brandriss, I. L. 1996. Writing in frost on a window pane: E-mail and chatting on RAM and copyright fixation. Journal of the Copyright Society 43:237-78.

Branscomb, A. W. 1995. Anonymity, autonomy, and accountability: Challenges to the first amendment in cyberspaces. Yale Law Journal 104:1639-79.

----- 1996. Cyberspaces: Familiar territory or lawless frontier, July 30, 1998. Journal of Computer-Mediated Communication 2(l): Part I. Available from the World Wide Web at: http://jcmc.huji.ac.il/vol2/issue1/.

Bruckman, A., compiler. 1997. The ethics of research in virtual communities. July 30, 1998. Available from the World Wide Web at: http://www.cc.gatech.edu/fac/asb/MediaMOO/ethics-symposium-97.htmI.

Cavazos, E., and G. Morin. 1994. Cyberspace and the law: Your rights and duties in the on-line world. Cambridge, MA: MIT Press.

Code of Federal Regulations. 1991. Title 45, Part 46 (Protection of Human Subjects). Washington, DC: Government Printing Office.

Commonwealth of Massachusetts (Information Technology Division, Legal Department). 1996. The basics of public key cryptography and digital signatures. August 4, 1998. Available from the World Wide Web at: http://www.magnet.state.ma.us/itd/Iegal/crypto-3.htm.

Curtis, P. 1992. Mudding: Social phenomena in text-based virtual realities. March 5, 1999. Available from the World Wide Web at: http://www.eff.org/pub/Net_culture/MOO_MUD_IRC/curtis_mudding.

Dctwiler, L. 1993. Identity, privacy, andanonymity on the Internet. July 30, 1998. Available from the World Wide Web at: http://www.eff.org/pub/Privacy/Anonymity/privacy-anonymity.faq.

Dibbell, J. 1998. My tiny life: Crime and passion in a virtual world New York: Henry Holt.

Donath, J. S. 1996. Identity and deception in the virtual community. August 4, 1998. Available from the World Wide Web at: http://www.judith.media.mit.edu/Judith/IdentityDeception.html

Electronic Frontiers Florida News. 1998. On-Line Copyright Infringement Liability Limitation Act. July 30, 1998. Available from the World Wide Web at: http://www.efflorida.org/Intell/hr3209.html.

Electronic Privacy Information Center Digital Signatures. August 4, 1998. Available from the World Wide Web at: http://www.epic.org/crypto/dss/.

Elkin-Koren, N. 1996. Public/private and copyright reform in cyberspace. Journal of Computer-Mediated Communication 2(2). July 30, 1998. Available from the World Wide Web at: http://jcmc.huji.ac.il/vol2/issue2/.

Ellison, C. M. 1996. Establishing identity without certification authorities. August 4, 1998. Available from the World Wide Web at: http://www.clark.net/pub/cme/usenix.html.

Froomkin, A. M. 1995. Anonymity and its enmities. Journal of On-line Law Art.4. March 5, 1999. Available from the World Wide Web at: http://www.wm.edu/law/publications/jol/froomkin.html.

GajjaIa, R. 1996. Cyborg diaspora and virtual imagined community: Studying SAWNET1. July 30.1998. Available from the World Wide Web at: http://ernie.bgsu.edu/~radhik/sanov.html.

Gray, M. 1996. Websites, hostnames and IP addresses, oh my. July 30.1998. Available from the World Wide Web at: http://www.mit.edu/-mkgray/net/terminoIogy.htmI.

Hardy, I. T. 1998. Project looking forward: Sketching the future of copyright in a networked world. August 5. Available from the World Wide Web at: http://lcweb.loc.gov/copyright/reports/.

Herring, S. 1996. Posting in a different voice: Gender and ethics in computer-mediated communication. In ***Philosophical perspectives on computer-mediated*** communication, edited by C. Ess, 115-45. Albany: State University of New York Press.

Information Infrastructure Task Force. ***1995. Intellectual property and the national information infrastructure: The report of the working group on intellectualproperty*** rights. Washington, DC: United States Patent and Trademark Office.

Jacobson, D. 1996. Contexts and cues in cyberspace: The pragmatics of naming in text-based virtual realities. ***Journal of Anthropological Research 52(4):*** 461-79.

King, S. A. 1996. Researching Internet communities: Proposed ethical guidelines for the reporting of results. ***The Information Society*** 12(2): 119-27.

Lea, M., and R. Spears. 1995. Love at first byte? Building personal relationships over computer networks. In ***Under-studied relationships: Off the beaten track,*** edited by J. T. Wood and S. Duck, 197-223. Thousand Oaks, CA: Sage.

Lee, G. B. 1996. Addressing anonymous messages in cyberspace. ***Journal of Computer-Mediated Communication*** 2(1): Part 1. July 30, 1998. Available from the World Wide Web at: http://jcmc.huji.ac.il/vol2/issuel/.

Long, G. P. 1994. Who are you?: Identity and anonymity in cyberspace. ***University of Pittsburgh*** Law ***Review*** 55: 1177-213.

Marvin, L. 1995. Spoof, spam, lurk and lag: The aesthetics oftext-based virtual realities. ***Journal of Cornpurer-Mediated Communication*** l(2). July 30.1998. Available from the World Wide Web at: http://www.ascusc.org/jcmc/vol1/issue2/marvin.html.

Mnookin, J. L. 1996. Virtual(ly) law: The emergence of law in lambdaMOO. July 30, 1998. ***Journal of Computer-Mediated Communication*** 2(l), Part 1. Available from the World Wide Web at: http://jcmc.huji.ac.il/vol2/issuel/.

Myers, D. 1987. “Anonymity is part of the magic”: Individual manipulation of computer-mediated communication contexts. ***Qualitative Sociology*** 10(3): 251-66.

National Institutes of Health, Office for Protection from Research Risks, Protecting Human Research Subjects. 1993. ***Institutional Review Boardguidebook.*** Washington, DC: Government Printing Office.

NUA Internet Surveys. 1998. ***How many on-line?*** July 30.1998. Available from the World Wide Web at: hnp://www.nua.ie/surveys/how_many_on-line/index.html.

Paccagnella, L. 1997. Getting the seats of your pants dirty: Strategies for ethnographic research on virtual communities. ***Journal of Computer-Mediated Communication 3(*** 1). August 5, 1998. Available from the World Wide Web at: http://www.ascusc.org/jcmc/vol3/issuel/paccagnella.html.

Parks, M. R., and K. Floyd. 1996. Making friends in cyberspace. ***Journal of Computer-Mediafed Communication*** l(4). July 30, 1998. Available from the World Wide Web at: http://www.ascusc.org/jcmc/vol1/issue4/parks.html.

Parks, M. R., and L. D. Roberts. 1997. Making MOOsic: The development of personal relationships on-line and a comparison to their off-line counterparts. July 30, 1998. Available from the World Wide Web at: http://psych.curtin.edu.au:80/people/robertsl/moosic.htm.

Post, D. G. 1996a. Encryption-It’s not just for spies anymore. July 30.1998. Available from the World Wide Web at: http://www.cli.org/DPost/X0009_ENCRYPTl.html.

——— 1996b. Knock knock, who’s there?: Anonymity and pseudonymity in cyberspace. July 30, 1998. Available from the World Wide Web at: http://www.cli.org/DPost/XOO12_KNOCK.html.

——— 1996c. Pooling intellectual capital: Thoughts on anonymity, pseudoanonymity, and limited liability in cyberspace. July 30, 1998. Available from the World Wide Web at: http://www.cli.org/DPost/paper8.html.

Quarterman, J. S. 1996. Summary: Third MIDS Internet Demographic Survey. Matrix News 6(3). July 30, 1998. Available from the World Wide Web at: http://www.mids.org/ids3/ids3sum.603.

Quarterman, J. S., and S. Carl-Mitchell. 1994. What is the Internet, anyway? ***Matrix News 4(8).*** July 30,1998. Available from the World Wide Web at: http://www.mids.org/what.html#The Matrix.

Rafaeli, S., F. Sudweeks, J. Konstan, and E. Mabry. 1998. ProjectH: A collaborative quantitative study of computer-mediated communication. In ***Networkand netplay: Virtual groups on the Internet,*** edited by F. Sudweeks, M. McLaughlin, and S. Rafaeli, 265-81. Cambridge, MA: MIT Press.

Rosenoer, J. 1997. ***Cyberlaw: The law of the Internet.*** New York/Berlin: Springer-Verlag.

Rosoff, D. 1995. ***PGP and what it does.*** July 30, 1998. Available from the World Wide Web at: http://www.arc.unm.edu/~drosoff/pgp/pgp.html.

Sewell, D. 1996. ***A taxonomy of usenet posting identities.*** July 30, 1998. Available from the World Wide Web at: http://packrat.aml.arizona.edu/-dsew/taxonomy.txt.

Spears, R., and M. Lea. 1994. Panacea or panopticon? The hidden power in computer-mediated communication, ***Communication Research*** 21(4): 427-59.

Thomas, J., ed. 1996. Introduction: A debate about the ethics of fair practices for collecting social science data in cyberspace. ***Information Society*** 12(2): 107-17.

Tien, L. 1996. Who's afraid of anonymous speech? McIntyre and the Internet. ***Oregon Law Review 75:*** 117-89.

Turkle, S. 1995. ***Life on the screen: Identity in the age of the Internet.*** New York: Simon & Schuster.

U.S. Copyright Office. 1996. ***Copyright basics.*** July 30, 1998. Available from the World Wide Web at: http://lcweb.loc.gov/copyright/circs/circ1.html.

——. 1998. On-Line Copyright Infringement Liability Limitation Act. August 5. Available from the World Wide Web at: http://lcweb.loc.goov/copyright/penleg.html.

Walther, J. B. 1996. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. ***Communication Research*** 23(l): 343.

Waskul, D., and M. Douglass. 1996. Considering the electronic participant: Some polemical observations on the ethics of on-line research. Information Society 12(2): 129-39.

Zakon, R. H. 1998. Hobbes' Internet timeline v.3.3. July 30, 1998. Available from the World Wide Web at: http://info.isoc.org/guest/zakon/Intemet/History/HIT.html.

DAVID JACOBSON is an associate professor of anthropology at Brandeis University. He has conducted fieldwork in Uganda, Grenada (WI.), Boston, Los Alamos, and cyberspace. His recent publicntions include Itinerant Townsmen: Friendship and Social Order in Urban Uganda ***(Waveland, 1986),*** Reading Ethnography ***(SUNY Press, 1991) and*** Spying Without Spies: Origins of America's Secret Nuclear Surveillance System ***(with Charles Ziegler, Praeger 1995). He has also published articles in the*** Journal of Divorce and Remarriage ***and the*** Journal of Anthropological Research.