The ethics of using the internet to collect qualitative research data

KAREN RODHAM and **JEFF GAVIN**

Department of Psychology, University of Bath, Bath, BA2 7AY, UK.

Email: psskr@bath.ac.uk

The practice of conducting research online is in its infancy. Consequently there is debate concerning the ethical implications of online data collection. We outline three approaches to online data collection and focus specifically on the issues of consent and anonymity of participants. We conclude that ethical issues raised when planning and implementing online data collection are no different to those raised by more traditional approaches to data collection.

Introduction

The internet is an invaluable resource for recruiting research participants, especially for studies that require recruitment of clinical or hard to find populations. Indeed, the prevalence of topic specific online message boards and e-mail lists, provides researchers with a ready-made sampling frame that they can utilize to access participants who are interested, affiliated and connected to the topic under investigation [1-3]. A further attraction is that vast amounts of qualitative data can be acquired with relative ease [4-6]. Collecting data in a textual form significantly reduces work efforts by eliminating the laborious task of transcription [1, 2, 7]. Although subtle, visual, non-verbal cues, which help to contextualize the interaction are lost [6, 7], it has been argued that what is effectively automatic transcription reduces human error, thereby enhancing the accuracy of data collection, and increasing the validity of any inferences made [4].

The practice of conducting research online is in its infancy and as a result, it presents dilemmas with which researchers must grapple [8]. Similarly, the formation of rules of conduct and ethical guidelines for the design and implementation of online research are at an early stage of development [9]. Consequently, researchers who wish to complete their research via the on-line environment, in the absence of existing accepted guidelines drawn up specifically for this kind of work, must think carefully about the ethical implications of their research. In the next section of this article, we outline three different approaches that we have employed to access qualitative data via the internet: online focus groups, online interviews and message boards, and highlight the ethical and methodological issues that employing these approaches can raise.

Assessment of qualitative data via the internet
Online focus groups
Online focus groups have featured with increasing prominence in the psychological literature, and have been used to explore a wide range of phenomena such as seasonal affective disorder [10], eating disorders [11] and racial violence [12]. Such groups can be defined as an interactive and intensive discussion between individuals guided by a specified research question [13]. Rather than conversing face-to-face, the accessibility of web-based 'host' programmes enable globally dispersed individuals to engage in a virtual discussion at their own convenience.

Consequently, vast amounts of information can be obtained with relative ease, in a short period of time [5], while retaining the essential quality of the focus group method: exploring the interaction between research participants enables the generation of a rich understanding of participants' experiences and beliefs [14].

Online interviews

Interviews can be conducted with respondents on a one-to-one basis through an exchange of e-mails. This type of interview can offer a suitable alternative for those respondents who require greater reassurance of confidentiality and anonymity, or who feel intimidated by the electronic presence of others in focus groups. The interview method also provides greater scope for in-depth probing of the respondents' experiences, thus enabling exploration of the complex feelings and attitudes embedded within an emotionally loaded topic.

Message boards

Message boards are areas on internet sites where individuals write messages to one another which can stimulate discussion, as other individuals read and respond to the messages. Researchers can access discussions that have been completed, or can follow discussions as they develop over time. Message boards can be 'open' where anyone can read, post or respond to messages or 'closed' where in order to read or participate in debates, individuals must register with the site and obtain a password to gain access.

On-line focus groups, on-line interviews and message boards can be used in a synchronous or real-time manner. In other words, researcher and respondents must be 'logged-on' at the same time, so that the questions which are posed get immediate responses, as in a face-to-face situation. However, these methods can also be used in an asynchronous manner. This presents numerous advantages over 'synchronous' or real-time methods. Being spread over time, asynchronous communication eases the pressure of immediate response, allowing the participant to respond at a time which is convenient to him or her, facilitating their reflexivity and deferring cognitive resources to the content rather than the management of the conversation [15].

Having outlined three key approaches to online data collection, we turn our attention to two ethical issues encountered when conducting online research: (1) gaining consent of participants and (2) confidentiality and anonymity of online participants. We discuss whether internet researchers face ethical dilemmas that are different to those faced by researchers employing other more traditional approaches.

Informed consent of participants

The British Psychological Society (BPS) states that *wherever possible* research investigators should inform all participants of the objectives of the research and of all the aspects of the research that may influence a participants' willingness to take part in the study [16]. We argue that the decision concerning how to address the issue of consent will differ depending on which approach to data collection via the internet is adopted. This decision will be informed by the debate about how public a forum the internet is. The level of 'public-ness' (both acknowledged and perceived) will vary for different environments and will therefore have different implications for the issue of gaining informed consent.

On-line interviews and focus groups

For approaches that require participants to actively engage with the researcher, either via on-line interviews or focus groups, we suggest that the general BPS guidelines should be followed whereby the researcher informs all participants of all aspects of the research so that they can decide whether or not they want to take part in the research. We also suggest that in these cases

the moderator of each website identified and selected through the screening process should be contacted and their consent sought for the proposed research. Ideally the moderator should be sent an introductory e-mail plus an outline of the proposed research and a summary of ethical considerations.

Once approval is gained, a participant recruitment notice can be posted on the website in order to recruit participants. This should outline the aims, purpose and nature of the study. Volunteers can then be asked to express their interest via e-mail. Those expressing a wish to participate should be promptly acknowledged and asked for confirmation of their age (if participants who are aged less than 16 years respond, consent should be sought from parents, whereas participants aged 16 years or above are deemed to be capable of giving their informed consent). Obtaining a person's informed consent means that the researcher has to be sure that the individual concerned is aware of the aims of the research that is being proposed, and is clear about what will happen to the research findings. The participant must also be made aware that they can withdraw from the research at any time. They will need to give their explicit consent to taking part in the research and give their permission for the researcher to record, analyze and report any data collected.

This process can be problematic because the internet allows individuals to conceal all or parts of their identity, or in fact allows them to adopt an alternative identity. Furthermore, and more worryingly, because the researcher is somewhat removed from the interactions, it is entirely possible that someone whom a researcher has excluded from the research because they are under 16 years of age, may at a later date, pose as someone over the age of 16, or may fake parental permission and therefore appear to be eligible to take part in the research. This problem was acknowledged in work with chronically ill children [16]. It was noted that there was no foolproof way of ensuring that the data collected had come from children, or indeed that the permissions for children to take part had come from their parents. This was acknowledged as a limitation of using the internet. However, this issue is not one which is exclusively a problem for those of us who gather our data via the internet. For example, researchers who collect data via means of questionnaires and interviews are also reliant on the participant to provide honest answers. It is no more difficult to create a different identity when completing questionnaires, answering an interviewer's questions or by taking part in online research. As we have mentioned, there is no guarantee that a questionnaire will be completed by the target person, or that interviewees or those taking part in focus groups 'offline' will express their 'true' feelings. Indeed, it may be that the internet is more suited to collecting such data as people feel freer to express their 'true' feelings. Thus, the validity of the responses elicited relies heavily on participants' integrity and honesty, regardless of the method used to elucidate the information.

Message boards

It is possible to argue that anyone who has access to the internet can view the content of 'open message boards' or post messages on them. Message boards could therefore be regarded as being in the public domain and consequently not subject to the requirement that the researcher obtains informed consent prior to using data collected from such environments (see BPS Guidelines). Individual contributions to the message board can therefore be considered in the same way as individual naturalistic observations in a public space. Furthermore, message boards are generally perceived and acknowledged by users as being public domains.

The 'open message board' is therefore considered to be both a public domain, as well as an environment where those posting and/or replying to posts would expect to be observed by others. Indeed, individuals posting onto an open message board are aware that their message can be read by anyone who has access to the internet. Indeed, it has been further argued that messages which are posted on such open forums are public acts, deliberately intended for public consumption, which means that researchers need not take more than 'normal precautions' when accessing such data [17]. We therefore argue that as long as researchers' maintain the confiden-

tiality of the individuals who have 'posted' and record data in a manner that would not cause personal identification, it is not necessary to seek consent of individuals using open message boards because ethical boundaries are not crossed.

In contrast, we suggest that obtaining and using data from closed message boards or chatrooms, without first gaining consent from individuals who use these media, would be unethical. These forums typically require those individuals who wish to use them to formally register with the site in order to gain access. Engagement with such an environment is therefore restricted only to those who have registered with the site and so can be considered to be both acknowledged and perceived as being more private than message boards. However, it is widely known that there are people who register with closed message boards in order to observe the discussions, rather than take part (such people are known as 'lurkers') [eg18]. It would be very straightforward for a researcher to do the same, in order access this data with relative ease, but we question the ethics of doing so in an environment that is deemed private or closed.

Thus, accessing such a chatroom or closed message board data without first obtaining the informed consent of those who are interacting in the environment which they consider to be relatively private, would be unethical. Furthermore, the BPS states that 'if the participants are likely to object or show unease once debriefed', withholding information is unacceptable. Accessing data in this way is therefore likely to be classed as a form of deception on the part of the researcher. It is therefore of paramount importance in this situation that the researcher reveals their status to the managers and users of these forums and seeks consent to use the text discussions in research.

Thus collecting data via the medium of the internet is subject to the same considerations as other methods of data collection. In other words the perceived and acknowledged privacy of the particular (in this case, internet) environment being targeted for research needs to be respected and the issue of when it is appropriate to seek informed consent carefully considered. In line with other commentators [19] we differentiate between data collected from private or semi-private sources (such as email or closed chatrooms) and those accessed through more open forums (such as bulletin or message boards).

Confidentiality and anonymity of participants

We have already mentioned that an important feature of the online research venue is the perception of anonymity it affords [20]. Anonymity has been shown to have a disinhibiting effect [1], social desirability is diminished [7] and the exchange of 'true' attitudes and opinions is encouraged [4]. This has been likened to the concept of 'strangers on a train', where a person feels able to disclose intimate details to a stranger sitting on the next seat [21]. The cost of divulging information via the internet (ie to a stranger) is significantly reduced [15], particularly when discussing sensitive topics.

Thus the anonymity of cyberspace allows internet users to express themselves in ways that may be constrained in their real world interactions. As recent identity-related research indicates, the internet is thought to enable the expression of a more 'authentic self' than is possible in daily life [20-22]. For some internet users the perceived anonymity of online activity affords an opportunity to give expression to normally withheld, or even pre-conscious, facets of themselves as they are freed from the inhibitions of face-to-face interaction. Indeed, it has been argued that the online persona is the 'core self', because as diminished social desirability reduces evaluative concerns, self-disclosure and the exchange of 'true' feelings are promoted [15].

However, it is important to point out that although online venues may appear private and encourage the disclosure of personal thoughts, they are, in reality public domains, that are relatively easy to access and afford little guarantee of absolute confidentiality or anonymity [2]. It is therefore important if engaging in the practice of 'harvesting' (collecting the words of others) from open sites, both the composers of the words and the name of the site hosting the message

board itself should be given pseudonyms in the write up of the research. Such practices serve to protect the anonymity of those whose words are being quoted and of the site concerned. This is even more salient when research is being conducted via more closed forums, and when asking participants for their consent, the 'public/private paradox' should be highlighted in order that participants can make an informed decision as to whether or not they wish to take part.

In addition, the public nature of the internet environment can make it impossible to forbid others who are not participating in the focus group research from reading participants' responses, thus effectively forming a breech of confidentiality. It is therefore important to ensure that potential participants are aware of the public nature of the forum before they consent to take part in the research. Linked to this, a difficult problem to resolve when conducting research in the internet environment is the inability to prevent non-participants from disrupting the focus groups through the posting of 'flaming' messages, whereby angry or insulting words about a person, issue or topic, are posted on the discussion forum. Whether this creates a differential effect both within and across the focus groups is hard to determine, although when it has happened in our research, the support that focus group participants offer one another (and the researcher) indicates that they are able to continue regardless of the outside intrusion.

It is possible with technical support, to set up closed on-line focus groups that participants who have expressed an interest in taking part in the research could sign up to. This would address the two problems of unwanted contributions from other internet users who are not signed up to the research and of non-participants being able to read the responses to questions posted during focus group discussions.

Conclusion

The practice of conducting research online is in its infancy and the formation of rules of conduct and ethical guidelines for the design and implementation of online research are at an early stage of development. Consequently, there is a lack of clear consensus about ethical issues such as confidentiality and consent. Most researchers who have written on the subject have distinguished between private or semi-private online venues and the more open forums, such as message boards. The general trend seems to be that the more private venues require more effort on the part of researchers to ensure that the individuals whose words are being used for research purposes are made aware of this, and their consent sought, whereas the more open forums can be considered to be public domains and treated as such in research terms, thereby the issue of consent is implied by the act of writing in such forums. Indeed, Mann and Stewart suggest that in sending or posting a message, there is an implied licence to read or even archive the information that the message contains [8]. However, there are some exceptions to this trend; Sharf for example, suggests that even when accessing open message boards, if a researcher plans to use a message, the person who posted it should be contacted and their explicit permission sought [9]. We suggest that this is an unnecessarily extreme position to take.

In this paper we have focused on two issues of primary ethical importance when designing and implementing research: gaining consent from, and anonymity and confidentiality of, participants. We believe that the decision-making process that internet researchers engage in when considering these issues should be no different to the process employed by researchers using more traditional means to access qualitative data.

We therefore argue that ethical issues encountered when planning and implementing research online are no different when conducting research by more traditional means. As psychologists, we find that the ethical code of conduct written by the British Psychological Society applies equally well to data collected via the internet as to other approaches to data collection. Providing the overarching principle of 'do no harm' is abided by, we argue that conducting research via the internet poses no more ethical dilemmas than when conducting research by more traditional means.

References

1. Griffiths M. Sex on the Internet: Observations and implications for Internet sex addiction. J Sex Res 2001; 38: 333-342.

2. Murray M, Fisher JD. The Internet: A virtually untapped tool for research: J Tech Hum Sci 2002; 19: 5-18.

3. Szabo A, Frenkl R, Caputo A. Deprivation feelings anxiety and commitment in various forms of physical activity: A cross-sectional study on the Internet. Psychologia 1996; 39: 223-230.

4. Franklin K, Lowry C. Computer mediated focus group sessions: naturalistic inquiry in a networked environment. Qual Res 2001; 1: 169-184.

5. Gaiser TJ. Conducting on-line focus groups. Soc Sci Comput Rev 1997; 15: 135-144.

6. O'Connor H, Madge C. Cyber-mothers: Online synchronous interviewing using conference software. Sociological Research Online 2001; 5(4).

7. Hewson C, Yule P, Laurent D, Vogel C. Internet research methods. London: Sage 2003.

8. Mann C, Stewart F. Internet communication and qualitative research: A handbook for researching online. London: Sage 2000.

9. Sharf B. Beyond netiquette: The ethics of doing naturalistic discourse research on the Internet. In Jones, S. (Ed) Doing Internet research: Critical issues and methods for examining the net. Thousand Oaks; CA: Sage 1999.

10. Michalak EE. The use of the Internet as a research tool: the nature and characteristics of seasonal affective disorder (SAD) amongst a population of users Interacting with Computers 1998; 9: 349-365.

11. Walstrom MK. "You know who's the thinnest?": Combating surveillance and creating safety in coping with eating disorders online. Cyberpsychology and Behaviour 2000; 3: 761-783.

12. Glaser J, Dixit J, Green DP. Studying hate crime with the Internet: What makes racists advocate social violence? J Soc Issues 2001; 58: 151-176.

13. Morgan DL. Focus groups. Annual Review of Sociology 1996; 22: 129-152.

14. Kitzinger J. The methodology of focus groups: the importance of interaction between research participants. Sociology of Health and Illness 1994; 16: 103-121.

15. Joinson AN. Understanding the psychology of Internet behaviour: Virtual worlds, real lives. New York: Palgrave MacMillan 2003.

16. British Psychological Society (2006). http://www.bps.org.uk/the-society/ethics-rules-charter-code-of-conduct/code-of-conduct/ethical-principles-for-conducting-research-with-human-participants.cfm (accessed Friday 2nd June 2006)

17. Fleitas J. Spinning tales form the world wide web: Qualitative research in an electronic environment. Qual Hlth Res 1998; 8: 283-292.

18. Pacagnella L. Strategies for ethnographic research on virtual communities. Journal of Computer Mediated Communication 1997; 3(1).

19. Winzelberg A. The analysis of an electronic support group for individuals with eating disorders. Computers in Human Behaviour 1997; 13: 393-407.

20. Waskul D, Douglass M. Cyberself: The emergence of self in on-line chat. Information Society 1997; 13: 375-397.

21. McKenna KYA, Bargh JA. Plan 9 from cyberspace: The implications of the Internet for personality and social psychology. Person Soc Psychol Bull 2000; 4: 57-75.

22. Rubin Z. Disclosing oneself to a stranger: Reciprocity and its limits. J Exp Soc Psychol 1975; 11: 233-260.

23. Adams J, Rodham K, Gavin J. Investigating the 'self' in deliberate self-harm, Qual Hlth Res 2005; 15: 1293-1309.

24. Turkle S. Life on the screen: Identity in the age of the Internet. New York: Simon and Schuster 1995.