Internet and Social Media Research: Additional IRB Guidance

Amanda Udis-Kessler, Permanent Chair of the CC IRB, September 3, 2021

This document is a supplement to the CC IRB webpage information about internet and social media research, focusing on some special challenges that this research can pose to the ability to guarantee that research is ethically and legally sound.

Some ideas in this document are based on a webinar titled "Hot Topics in Online Survey Research: Participant Identification, Consent, and Risk," which was held by Public Responsibility in Medicine and Research (PRIM&R) on March 12, 2015. While this document has not been reviewed by PRIM&R or the presenters (Elizabeth Buchanan, PhD, University of Wisconsin-Stout, and B.R. Simon Rosser, PhD, MPH, LP, University of Minnesota School of Public Health), both parties have graciously given permission for materials from the webinar to be referenced in the creation of this document.

1. Risk-Benefit Ratio

If participants are not who they claim to be or if the same individual repeats a study multiple times in order to make more money from it or just to throw a monkey wrench into the study, a research project loses its claim to validity (the study showing what it claims to show) and reliability (the study can be replicated by other researchers), and thus loses its benefit to academia or society. If research findings do not benefit academia or society a project fails the risk/benefit calculus and cannot be approved by an IRB. Thus, the question of participant identity verification is an IRB matter in the sense that the IRB must be convinced that a researcher carrying out internet research has addressed this issue.

Several possible strategies have been suggested by researchers and IRB experts:

- Find a way to provide unique personal ID numbers to participants in order to verify that people are who they say they are
- Use monitoring software such as SafeSurf to screen out children as participants if the research is aimed at adults
- If acceptable (a study is minimal risk and anonymity is not required) consider requiring ID information from participants, possibly including partial credit card information or social security numbers
- In some cases, in-person or Skype recruitment may be possible, and can be combined with ID information collection
- Use check boxes in the consent form so that participants must affirm (by checking the box) that (for example) they are male (if the study is only for men) or over 17 (in almost all cases); someone still may check a box inappropriately but some potentially inappropriate participants may choose not to participate rather than flagrantly lying
- Consider de-duplication (flagging responses from related IP addresses and payments requested to similar names or PayPal emails, using software followed up with manual checks of suspicious responses)
- Collect and compare IP addresses (both entire addresses and address sections), bearing in mind that this strategy is not perfect (someone at an internet café could complete the same

survey multiple times from different computers); this strategy should only be used if participants are informed during the informed consent process that IP addresses will be collected and checked, and it is not appropriate for sensitive research involving more than minimal risk of harm
- Carry out cross-validation by checking consistency in answers at the beginning and end of a survey and possibly during the survey as well; items to be checked might include age, zip code, or important demographic traits tied to the survey's research goals
- Qualtrics enables researchers to check a box that makes it harder for automated "bot" trolls to find and participate in studies; this box should generally be checked unless there is a very important reason to leave it unchecked

The CC IRB will work with researchers to help them think through the appropriateness of these various options for the research project at hand.

2. Recruitment Methods and Advertising Material

The internet may be used to advertise both internet and other types of research in a wide range of ways, including online advertising, chatroom postings, Twitter streams, blog postings, YouTube videos, email solicitations, texts, and (undoubtedly over time) other options as well. All forms of recruitment and advertising for research are equivalent in the IRB requirement that they must be mentioned in the IRB application and reviewed by the IRB. Traditionally, it has been easy to send a draft of a flyer announcement or an email as part of the application process. As researchers use a wider variety of online recruitment methods they will need to find ways to show the materials to the IRB (for example, a link to a YouTube video embedded in an IRB application). If researchers have any questions about how to provide online recruitment materials to the CC IRB they should check with the Chair.

3. Informed Consent Process and Documentation

All research that requires IRB review must include an appropriate informed consent process and appropriate documentation of consent to receive IRB approval. Some aspects of internet research consent may be similar to face-to-face research, and in some cases the differences are relatively minor. For example, all appropriate consent forms include information about the length of time participation in the overall project is expected to take for any one participant; an online modification of this element might include mention of whether a particular internet study requires more than a single period at the computer (surveys with follow-up components, for example).

That said, there are some important differences between online and face-to-face research regarding potential confidentiality and privacy issues. For example, if a researcher has determined that IP addresses need to be collected and stored for participant identity verification purposes the researcher must include that information in the consent documents and may not treat a project as anonymous in that case. Moreover, participants have a right to know about aspects of the research such as how data will be transmitted, whether the research project is using a survey host such as SurveyMonkey or Qualtrics, whether their information will be encrypted,

and whether the survey host (if one is used) will retain identifiable information, among other information related to potential confidentiality/privacy issues.

Researchers should include language in the consent documents taking care to inform participants about potential limitations to confidentiality and privacy, such as the following:

- "Although every reasonable effort has been taken, confidentiality during actual internet communication procedures cannot be guaranteed due to the possibility of interception of data" (modified from Penn State University IRB website)
- "Data may exist on backups or server logs beyond the timeframe of this research project." (from University of California, Berkeley IRB presentation)
- "Please note that the online survey is hosted by [Qualtrics, SurveyMonkey, etc.] which is a web survey company located in the USA. All responses to the survey will be stored and accessed in the USA. This company is participant to US laws, in particular, to the US Patriot Act/Domestic Security Enhancement Act that allows authorities access to the records including your answers to the questions. The security and privacy policy for [Qualtrics, SurveyMonkey, etc.] can be viewed at http://..." (modified from Prim&r presentation, "Ethical Internet Research: Informed Consent Regulations and Realities")

Another type of online research, the study of online communities, raises both privacy and consent issues because the CC IRB defines sites that require users to create accounts (including a login and a password) as private spaces even though in theory anyone may create such an account and join the community. This issue is even more important in online communities built around identities that may involve social stigma. Therefore, the CC IRB requires that researchers recruiting in, or gaining information from, online communities have a rigorous consent process in which all parties present in the community at the time of research must consent to the researcher's presence, observations (even if these only entail lurking), and interactions with the community. This is likely to mean requesting consent every time the researcher logs in and may involve consent requests made to every new person as they sign in. As with other circumstances where either the consent process or consent documentation can make research impracticable researchers may petition the IRB for a waiver of consent or documentation, but a researcher would need to provide strong evidence that such a waiver would not lead to harm in order for the IRB to approve the waiver for this type of research.

While internet research poses challenges on certain fronts, it may sometimes call for a consent process that is actually simpler than the standard consent form common in face-to-face interviews and experiments. Most survey research, and notably most survey research carried out by Colorado College researchers in recent years, qualifies as minimal risk, especially if minimal risk is redefined somewhat to mirror everyday experiences on the internet. In most cases, if someone using the internet encounters an image or phrase that is disturbing, the person simply closes the browser window. Survey research that is no more disturbing than what someone may encounter surfing the web can arguably be seen as minimal risk. (Language addressing this issue in a consent form might read as follows: "We anticipate that your participation in this study presents no greater risk than everyday use of the internet.")

If most survey research involving adults is minimal risk it becomes important to use a consent process that will not cause most respondents to drop out before getting through the process. This

priority is in part a matter of the risk/benefit ratio; studies that require a certain number of participants improve their risk/benefit ratio by making the consent process straightforward enough to keep a sufficient number of potential participants committed to participating. Minimal-risk internet survey projects that call for a substantial number of participants may therefore receive IRB permission to modify the consent process to make the research more practicable. A related point is that certain types of potential participants, such as racial minorities and the poor, are more likely to drop out if the consent process is experienced as too burdensome, and the Belmont principle of justice directs researchers to work not to exclude demographic groups when possible.

Recent research on simplifying consent processes for internet research has indicated that the following strategies might be used for minimal-risk surveys when documentation of informed consent can be waived:

- Design a consent process with more screens while including less material on each screen, leaving sufficient "white space" and using sufficiently large fonts for the text to be easily readable (a potential participant reads one screen, clicks a consent button related to the material on that screen, and is taken to the next screen)
- Design consent buttons that include the specifics of what's being consented to on the button itself (for example, "I consent to answer questions about my experience with_____")
- Use check boxes and request that participants check off different aspects of the study to indicate that they understand the nature and potential consequences of the research; an example might be, "I understand that some of the questions might cause me to have an emotional response"
- Design two versions of the consent form, an abbreviated version and a complete version. The abbreviated form would include information that virtually all participants would want to know, such as who's carrying out the research, the purpose of the research, the number of survey questions, topics covered, the probable length of time to complete the survey, any payment or incentives, and the minimum age one must be to participate. This basic information would be followed by two options: a "start survey" button and a "more information" button. Participants clicking on the "start survey" button would be acknowledging that they have as much information as they feel they need. Participants clicking on the "more information" button would be directed to the complete consent form, at the end of which there would be a "start survey" button that would lead to the instrument.

One other less obvious element of consent needs mentioning here: in a paper survey a participant is always free to skip an item. Some online surveys are designed so that any question left unanswered prevents the participant from continuing on. In order to avoid coercion (and also to improve response rates), all online survey questions should include a "decline to answer" option that, if checked, moves the participant forward to the next question or screen.

4. Deception and Incomplete Disclosure

One clear advantage of face-to-face research involving deception is that it is usually fairly straightforward to include a good debriefing process at the end of the research experience. Internet research involving deception is more challenging in that there is no way to guarantee that a participant will not close out of their browser window before reaching the debriefing page.

When possible, participants should indicate that more information about the study will provided at the conclusion of participation; a participant who completes the study will know to look for a final page with debriefing information. However, some deception studies (particularly certain studies in social psychology and marketing) cannot indicate the possibility of deception during the consent process for reasons described earlier in this document. In such cases, the researcher must assume that at least some participants will never see the debriefing page and should design their projects accordingly.

The most important component of a deception project in which participants may terminate participation before debriefing is that the deception aspect of the project must be minimal-risk or harmless. For example, social psychology studies involving a distraction element (a task falling between the two tasks of interest to the researcher) in which the participant does not know that the distraction element is serving in this way would not be considered harmful even though the participant does not have complete information about what they are doing. Even deception in which a participant is intentionally misled may be minimal-risk if the participant is given misinformation about something that cannot cause emotional, physical, financial, academic, legal, criminal, or social harm. Again, social psychology offers useful examples. If a participant engages in a harmless online activity and is told that the activity measures one thing when in fact it measures something else it may not be ethically problematic if the participant drops out part-way through and never learns what the activity was actually measuring.

Online research involving any sort of deception will receive substantial IRB scrutiny but when such research is harmless it is much more likely to be approved by the IRB in an expedited process.

5. Privacy and Confidentiality

Confidentiality and privacy face special challenges in online research. Data security issues play an important role in these challenges, and it is the responsibility of researchers to both understand various security concerns and explain them to participants during the consent process:

- The survey software being used and whether it is officially licensed through Colorado College (Qualtrics) or from a third party (e.g., SurveyMonkey)
- The types of confidential information being collected (e.g., IP address, email address)
- Security measures during data transmission from browser to server as well as for data stored on a server
- What information will be stored and how and where it will be stored
- How identifying information will be de-linked from survey data
- How long log files are kept
- Whether data are date and time stamped
- How the particular platform handles data back-ups
- The privacy and confidentiality policies of the survey software company; and
- What happens to any copies of the data files once the research is complete

Encryption should be used both for data transmission (Secure Socket Layer (SSL) protocol) and for data storage. When encryption can be used the need for warnings about possible breach of confidentiality and privacy may be mitigated, though some warning is always appropriate.

It is the researcher's responsibility to be aware of any security concerns that have arisen regarding the software or platform being used. For example, a 2013 research study found that Amazon Mechanical Turk (mTurk) has a data security vulnerability that can allow mTurk worker IDs to be connected to personally identifying information that mTurk workers post on their Amazon profile pages. Researchers using mTurk must determine whether this security concern has been addressed successfully, and if not, must include a warning about the concern in consent documents. (Contact the IRB Chair for more information about this 2013 study.)

Issues of privacy may also be complicated in research involving online communities, since members of such communities usually assume that what they say to each other remains within the community. Online communities blur traditional boundaries of public and private in certain ways and so, to maximize the ethical nature of research, should generally be treated as private spaces. Comments quoted from an online support group in published research may be traceable back to a particular individual (whether their internet persona, such as an avatar, or their real-world identity), potentially making the group unsafe for the person.

One way to address this concern is for the researcher to individually contact any participant the researcher wishes to quote and to receive direct permission for the use of a specific quote. Similarly, names of avatars and other internet personas may only be used with the participant's consent, just as would be true of their real-world name. Such individual contact can usually take place through a private messaging system such as Facebook PM or GoogleChat.

6. Vulnerable Participants

Internet research involving children is generally discouraged by the CC IRB. Beyond the issues discussed above, the question of how to obtain and to be sure that one has obtained parental permission remains thorny. An internet researcher cannot be absolutely sure they are interacting with a child, or, for that matter, with the child's parent. Moreover, in addition to the legal protections discussed elsewhere in this document, on which most IRB policies are based, researchers are legally beholden to the Children's Online Privacy Protection Act (COPPA; information at http://www.ftc.gov/ogc/coppa1.htm). COPPA's focus is children under 13 years old, and it requires website operators to post a privacy policy on their website and create a mechanism by which parents can control what information is collected from their children and how such information may be used.

Seventeen-year-old Colorado College students are also legally children; the IRB has consulted with Colorado lawyers and determined that even a 17-year-old college student is not legally authorized to enter into a binding contract, which includes the signing of a consent form. Internet research involving Colorado College students as participants must include language in the consent form or paragraph in which, by agreeing to participate in the study, the student confirms that he or she is 18 years of age or older. Consent forms without such language will not be approved by the CC IRB.