Security Research with Human Subjects: Informed Consent, Risk, and Benefits

Maritza Johnson, Steven M. Bellovin, and Angelos D. Keromytis

Columbia University, Computer Science Department
{maritzaj,smb,angelos}@cs.columbia.edu

Abstract. Computer security research is facing a growing trend of researchers collecting data directly from users or their personal devices. Many researchers are required by law to obtain the approval of an ethics committee for research with human subjects. This process is designed to ensure the ethical treatment of subjects and focuses on key concepts such as informed consent, minimized risk, and generally maximizing benefits to the research subjects. Computer security researchers who conduct human subjects research should be concerned with these aspects of their methodology regardless of whether they are required to by law, it is our ethical responsibility as professionals in this field. Previous discourse on the ethics of computer security research fails to satisfactorily address how the nature of security research may complicate the process of determining how to treat human subjects ethically. We suggest that our community take an active role in crafting best practices for how to treat human subjects ethically.

Key words: security research, human subjects, responsible conduct, ethics review committee, institutional review board

1 Introduction

Computer security researchers have found that in many cases the use of real data is critical to forming an accurate model of the problem and is the best way to evaluate the effectiveness of a proposed solution. For many of these problems obtaining real data means going directly to actual users. Data collection could require installing monitoring software on a user's personal machine, instrumenting a website, or conducting a laboratory study. As research involving human subjects becomes more prevalent it becomes increasingly important for us to consider the relationship between the researchers and the subjects.

At many institutions an ethics committee must approve the research methodology before human subjects research can be conducted. The committee's primary task is to ensure participants are treated ethically. The review includes a focus on concepts like informed consent, minimized and acceptable risk, respect for persons, and beneficence. As computer security professionals, regardless of whether a review is required by law, we should determine how to address these concepts by discussing what it means to treat participants ethically in our field.

The existing regulations were written for biomedical and behavioral research when a few researchers demonstrated the need for an ethics review [15] [9]. Thus, a fruitful discussion includes determining how our research methodologies are similar and different from these fields.

In the context of computer security research ethics, this paper is concerned with the ethical treatment of human subjects. Prior work has mentioned this topic as a subproblem related to the larger concern that is how to conduct security research ethically, or it has focused on specific examples. One of the earlier papers addressed how to ethically design phishing experiments [7], it included detailed advice on how to work closely with an IRB to assuage their concerns, and how to navigate the process of waiving informed consent. Later it was suggested that security researchers should familiarize themselves with cases where IRB review is required [8]. Allman noted the need for community ethical standards to alleviate the burden program committees face when reviewing an ethically questionable paper [2]. Cranor enumerated specific examples of how security and privacy human subjects research may introduce additional risk to participants [4]. Matwyshyn *et al.* suggested the community establish best practices for doing vulnerability research [14]. Kenneally *et al.* introduced the Ethical Impact Assessment (EIA) framework to guide the process of determining the potential risks and benefits for stakeholders [11].

The EIA framework is a useful starting point for bringing concepts like informed consent and beneficence to the attention of researchers, however the framework does not capture the nuances that make computer security human subjects research subtly different from other types of human subjects research. The EIA framework could benefit from the anecdotal experience of designing human subject experiments. With this in mind we focus on why computer security research should continue to discuss the ethics of human subjects research, starting with a discussion of what makes our research different and similar to other fields. We also address the likelihood that IRBs of U.S. universities are prepared to evaluate our research methodologies.

2 Computer Security Research with Human Subjects

To begin the discussion of computer security research with human subjects we compare and contrast our field with the two primary fields of human subjects research, biomedical and behavioral research. Our community should engage in an active discussion of how to apply the concepts of informed consent, risks, and benefits. Examples of how security human subjects research differs from the usual behavioral science research include: detailed collection of potentially sensitive data, observation of login credentials, studies that involve actively attacking the subject, and the need to hide the true purpose of the study [4]. Ethics committees are expected to determine whether subjects are properly informed and consent to participation, and that potential risks are minimized while benefits are maximized when possible. What do these terms mean in regard to computer security research: informed consent, risks, and benefits? How can security researchers ensure that they have appropriately addressed each of these?

Behavioral science researchers have a history of attempting to distinguish themselves from biomedical research, where in biomedical research the risks tend to be of a physical nature [13]. In 2003 a more comprehensive list of possible harms was published: physical, psychological, social, economic, legal, and dignitary [3]. Computer security research is more like behavioral research where the potential risks tend to be of categories other than physical, these risks can be difficult to quantify and difficult to describe.

Informed consent has two aspects, the first is that the participant is presented with the potential risks and benefits of participation, the second is that the participant decides whether to participate free of coercive pressures. The first aspect may be of concern for our field. Empirical evidence has shown that the typical user has an inaccurate mental model of how basic computer security primitives operate [18], and how common attacks like phishing are executed [5]. If they are asked to install monitoring software on their personal device, how can they be expected to properly evaluate the risks of participating unless the potential risk is very clearly explained in layman's terms? The IRB process evaluates whether the consent form is understandable to potential subjects. We question whether written consent forms are effective. Researchers from other fields have attempted to evaluate the effectiveness of various mediums [1]. In some cases it may be useful to engage in a conversation where the researcher explains key ideas and the participant is given a chance to ask questions. For some studies this is unrealistic but applying it when possible could be beneficial.

In some cases disclosing the research purpose in the consent form may threaten the validity of the results. If a researcher plans to study how users respond to a specific type of attack, or measure a user's security or privacy mindedness, revealing the purpose of the study will affect the participant's behavior. To avoid this researchers can request a waiver of informed consent, or they can mask the true purpose of the study. Obtaining a waiver typically requires demonstrating that the potential risks are minimal and that other study designs will not suffice. In many cases the IRB will request that participants are debriefed once the study is completed, this can serve as a tool to reduce the perceived risks.

Debriefing subjects may be a point where our field differs from behavioral research. Sometimes revealing the true research protocol may cause the participant distress. For example, Milgram debriefed his participants and revealed that the study was actually about obedience [15]. It would not be surprising if participants were embarrassed or otherwise disturbed by their own behavior. In our field, however, debriefing can be an opportune time to increase the benefits of participation.

At the most basic level debriefing a participant after the study gives the researcher a chance to address the concerns of participants or explain the purpose of the study. For studies where participants are being attacked or are answering questions related to their security knowledge and practices, the debriefing period is also an opportunity to educate users. However, it can be difficult to design an effective debriefing message, especially when users participate remotely and are not present in person. Depending on the research topic, the researcher may be in the position to give advice that is known to be effective [12], or they may feel debriefing will raise more concern than it is able to address thus causing unnecessary harm to participants [7]. It would be useful to have guidelines to help a researcher decide when each technique is appropriate or desirable, perhaps it depends on the amount of risk involved.

2.1 Scenarios

Here we present scenarios of computer security human subjects research and discuss difficulties that may arise in evaluating informed consent, risks, and benefits.

Online Social Networking In our own research, we are interested in the privacy settings of online social networking users. Using the Facebook API we developed an application that accesses data associated with the user's profile and ask questions about specific items. This scenario is an example of a study where it is necessary to access a large amount of private data, not only the participant's private data but also data associated with friend profiles. Depending on how active a user is, and how long they have been a member of the social networking site, the user could have a large amount of data associated with their profile. Approximately 5 of about 300 participants commented on the application's request to access all profile data, other participants either did not notice or perhaps do not care that this was a requirement. The description of the research carefully describes how information is accessed and what information is stored, namely only the participant's responses to our questions, not the actual data. We wondered whether participants read the consent form or merely clicked agree as users are known to do with EULAs [10].

As part of the study we alert users of inconsistencies between their sharing intentions and their settings[1]. Ideally, we would be able to directly modify their privacy settings as a benefit, however this is not possible through the API and it is unlikely that we could craft their ideal policy based only on these inconsistencies. The participants who have many inconsistencies between their intentions and settings obviously need help with configuring their privacy settings but we are not in the position to give them practical advice. Thus, we're able to identify participants who need assistance but at this stage of the research cannot offer real help.

Monitoring Software Installation A security researcher plans to collect data by installing monitoring software on personal devices. The software would collect and report data to the researchers once per day. This scenario raises several interesting ethical questions. In regard to informed consent, is the participant

[1] Columbia University Protocol IRB-AAAF1543

aware of the capabilities of the monitoring software installed? Do they understand the implications? Is there anyway for them to verify that the researcher is only collecting the data that was agreed upon?

If the researcher is collecting data about a specific exploit and notices behavior that may be from a different exploit, what is the proper course of action? The researcher wants to collect data about this new exploit and they may not want to halt data collection as they wait to get approval from the participant or an ethics committee. The risks to the participant include being unaware that this additional data is collected. If the data led to a patch for the exploit the participant could benefit, but there may be a considerable delay before this would materialize. How immediate must a benefit be in order for it to be a benefit of participation?

Publicly Available Password File Consider a password file that has been posted on the web. This data is publicly available on the Internet but it was stolen by someone other than the owner of the website [17]. Since the data is now publicly available is it ethical to use it in research? The data are valuable to those who are doing research on real-world password choices, but it is acknowledged that the files are stolen. In addition, there is personally identifiable information in some of the files, where email addresses were used as a password. In the past there were ongoing debates on whether unethically obtained data should be used by researchers [16].

3 Institutional Review Boards

In many regions human subjects research requires the approval of an ethics committee (E.U.) or an IRB (U.S.). Their primary task is to ensure the ethical treatment of human subjects in proposed research. It has been suggested that an IRB may be in the best position to evaluate the ethics of research methodologies, as opposed to a conference program committee [2]. This raises the question – are IRBs prepared to evaluate computer security protocols? Some have suggested perhaps they are not, though no data have been presented [8] [6].

The federal regulations were written with biomedical and behavioral research in mind. As discussed in the previous sections, while computer security research shares similarities with these fields, there are unique qualities that suggest the regulations established may need to be reevaluated in regard to computer security. Determining the veracity of this statement is a topic for future research. As a first step we suggest looking at the composition of existing IRBs.

The U.S. Department of Health and Human Services (HHS) specifies membership requirements and criteria to ensure the board represents a range of backgrounds and experience.[2] The regulation suggests that the IRB should select members based on the types of research an institution typically conducts, providing the best coverage for the common case. WIth the increase in computer

[2] "IRB Membership, Protection of Human Subjects." 45 CFR 46.107, 2009.

security research involving human subjects, we wondered if IRB membership had been affected. We searched the IRB websites of forty U.S. universities with computer science departments.[3] We sought to answer the following questions:

- Is the IRB membership list available online?
- Does the board have a computer science or computer engineering member?

HHS requires IRBs to submit an updated membership roster each year that includes each member's name and their academic department affiliation. We hypothesized that many IRBs would make this list publicly available on their website since they are required to keep an updated roster on record. We found that 17 of the 40 universities post their membership roster on their website. This number is significantly lower than expected.

We hypothesized that few schools would have a computer science member. We found that 5 of the 17 IRBs that post their roster online have a member with a computer science or computer engineering background. This number is slightly higher than expected, though only one of the five is a security researcher (based on information available on member home pages). All 5 represent universities with a sizable human-computer interaction (HCI) program, this makes sense considering HCI entails human subjects research.

We include a table of the forty universities, indicating if the roster was online, if they have a CS/CE board member, and whether they have separate boards for medical and non-medical research (Appendix A). The presence of a non-medical board is of interest because it demonstrates that behavioral researchers have successfully made a case that their research differs from biomedical research, in some cases it just indicates the absence of a medical school.

4 Conclusion

In this paper we discuss characteristics of computer security human subjects research that suggest we should focus on the concepts of informed consent, risks, and benefits as they pertain to our research. Concrete suggestions for how to approach the task are noticeably missing from our discussion, more work needs to be done before the issues can be fully described. We should continue to explore best practices for our field to ensure the ethical design of research methodologies.

We also present results from a small survey of university IRBs to begin the discussion of the adequacy of existing boards. A more thorough study would interview a larger set of IRBs and perhaps inquire about their experiences to date with computer security research. The interviewer could focus on measuring the number of protocols the CS department submits each year, when the first protocol was submitted by the CS department, and whether they have had any concerns specific to computer security or their level of technological experience. If such a study found that IRBs have experienced problems evaluating computer security protocols members of our community could consider offering their assistance to IRBs as an external source of expertise when necessary.

[3] The universities were chosen based on the 2010 U.S. News & World Report rankings of computer science graduate schools.

Acknowledgment

The authors would like to thank the anonymous reviewers for their help in improving the submission. This work was supported by the NSF through Grant CNS-09-14845. Any opinions, findings, conclusions or recommendations expressed herein are those of the authors, and do not necessarily reflect those of the US Government or the NSF.

References

1. Patricia Agre, Frances A. Campbell, and Barbara D. Goldman *et. al.* Improving informed consent: The medium is not the message. *IRB: Ethics and Human Research*, 25(5):S11 – S19, Sep. – Oct. 2003.
2. Mark Allman. What ought a program committee to do? In *WOWCS'08: Proceedings of the conference on Organizing Workshops, Conferences, and Symposia for Computer Systems*, pages 1–5, Berkeley, CA, USA, 2008. USENIX Association.
3. National Research Council. *Protecting Participants and Facilitating Social and Behavioral Sciences Research.* National Academies Press, Washington D.C., 2003.
4. Lorrie Cranor. Ethical concerns in computer security and privacy research involving human subjects. In *Financial Cryptography and Data Security*, Lecture Notes in Computer Science, pages 247–249. Springer Berlin / Heidelberg, 2010.
5. Rachna Dhamija, J. D. Tygar, and Marti Hearst. Why phishing works. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 581–590, NY, NY, USA, 2006. ACM Press.
6. David Dittrich, Michael Bailey, and Sven Dietrich. Have we Crossed the Line? The Growing Ethical Debate in Modern Computer Security Research. In *(Poster at) Proceedings of the 16th ACM Conference on Computer and Communication Security (CCS '09)*, Chicago, Illinois, USA, November 2009.
7. Peter Finn and Markus Jakobsson. Designing and conducting phishing experiments. In *IEEE Technology and Society Magazine, Special Issue on Usability and Security*, 2007.
8. Simson L. Garfinkel. IRBs and security research: myths, facts and mission creep. In *UPSEC'08: Proceedings of the 1st Conference on Usability, Psychology, and Security*, pages 1–5, Berkeley, CA, USA, 2008. USENIX Association.
9. James H. Jones. *Bad blood: The Tuskegee syphilis experiment.* Free Press, New York, 1981.
10. Matthew Kay and Michael Terry. Textured agreements: re-envisioning electronic consent. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, SOUPS '10, pages 1–13, NY, NY, USA, 2010. ACM.
11. Erin Kenneally, Michael Bailey, and Douglas Maughan. A framework for understanding and applying ethical principles in network and security research. In *Financial Cryptography and Data Security*, volume 6054 of *Lecture Notes in Computer Science*, pages 240–246. Springer Berlin / Heidelberg, 2010.
12. Ponnurangam Kumaraguru, Justin Cranshaw, Alessandro Acquisti, Lorrie Cranor, Jason Hong, Mary Ann Blair, and Theodore Pham. School of phish: a real-world evaluation of anti-phishing training. In *SOUPS '09: Proceedings of the 5th Symposium on Usable Privacy and Security*, pages 1–12, NY, NY, USA, 2009. ACM.

13. Susan M. Labott and Timothy P. Johnson. Psychological and social risks of behavioral research. *IRB: Ethics and Human Research*, 26(3):11–15, May – June 2004.
14. Andrea M. Matwyshyn, Ang Cui, Angelos D. Keromytis, and Salvatore J. Stolfo. Ethics in security vulnerability research. *IEEE Security and Privacy*, pages 67–72, 2010.
15. Stanley Milgram. Behavioral study of obedience. In *Journal of Abnormal and Social Psychology*, pages 371–378, 1991.
16. Stephen G. Post. The echo of Nuremberg: Nazi data and ethics. In *Journal of medical ethics*, pages 42–44, 1991.
17. Skull Security. Passwords - skull security. `http://www.skullsecurity.org/wiki/index.php/Passwords`.
18. Alma Whitten and J. D. Tygar. Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In *8th USENIX Security Symposium*, pages 169–184, 1999.

Table 1. Appendix A: Institutional Review Board Data

University	Roster Online	CS/CE Member	Non-medical Board
CMU	N		
MIT	Y	Y	N
Stanford	Y	N	Y
Berkeley	Y	Y	Y
Cornell	Y	N	N
Univ. of Illinois Urbana Champaign	N		
Univ. of Washington	Y	N	
Princeton	Y	N	N
Univ. Texas	Y	N	N
Georgia Tech	Y	Y	N
Cal Tech	N		
Univ. of Wisconsin Madison	N		N
Univ. of Michigan	Y	Y	N
UC San Diego	N		
UC Los Angeles	Y	N	Y
Univ. of Maryland	N		
Columbia	Y	N	Y
Harvard	N		Y
Univ. of Pennsylvania	N		Y
Brown	N		
Purdue	Y	N	Y
Rice	N		
Univ. Mass	N		
UNC Chapel Hill	N		
Univ. Southern California	N		Y
Yale	Y	N	
Duke	N		
Johns Hopkins	N		Y
New York Univ.	N		Y
Ohio State	Y	N	
Penn State	N		Y
Rutgers	N		
UC Irvine	Y	Y	Y
Univ. of Virginia	N		
Northwestern	Y	N	Y
UC Santa Barbara	N		
Univ. of Chicago	Y	N	Y
Univ. of Minnesota	N		
UC Davis	N		Y
Univ. of Colorado Boulder	N		