Tell Me Lies: A Methodology for Scientifically Rigorous Security User Studies

Serge Egelman
Brown University
Providence, RI 02912
egelman@cs.brown.edu

Janice Y. Tsai
California Council on Science & Technology
Sacramento, CA 95814
janice.tsai@ccst.us

Lorrie F. Cranor
Carnegie Mellon University
Pittsburgh, PA 15213
lorrie@cs.cmu.edu

CHI 2010, April 10–15, 2010. Atlanta, Georgia, USA
ACM 978-1-60558-930-5/10/04.

Abstract

Studies that examine users' perceptions of online privacy and security are especially difficult to design because the study participant must be in a similar mindset as they would be in real life. To test a system designed to protect users from a known risk, study participants must be made to believe that they are actually at risk, otherwise their resulting behaviors cannot be generalized to the real world. At the same time, ethics issues arise when the study participant is actually placed at risk. In this paper, we describe our methodologies when performing usable security experiments, and we argue that deception is a necessary component when performing human subjects experiments in the areas of privacy and security.

Keywords

Security, privacy, usability, study methodologies, scientific validity

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User interfaces – Evaluation/ methodology, H.5.2 User Interfaces: User-centered design, H.5.3 Group and Organization Interfaces: Evaluation/methodology, collaborative computing; K.4.1 Public Policy Issues: Privacy.

Introduction

The results of an experiment are only as valid as the methodology used to design said experiment. The study of human factors in online privacy and security is an area that requires even more attention to detail when designing user studies because data on the user's primary task is rarely the objective. That is, users do not usually sit down at the computer to "do security;" security is often seen as an impediment to completing another task, and it is not a task unto itself [5]. Usability studies that frame security as the primary task are often flawed because their results cannot be generalized to users' behavior in their natural environments. As an additional constraint, if study participants are aware that online security behaviors are being studied, they may alter their actions to "succeed" in the study. Thus, to yield scientifically valid results from user studies relating to online security and privacy, we are forced to deceive our participants about the nature of our studies.

The measures that people take to increase their online privacy and security often come at a cost—time or money—and, therefore, a rational person would only take these measures voluntarily when she believes she is legitimately at risk. For instance, it is not rational for a user to create a long unmemorizable password to protect information that is already public, because nothing is at stake. Similarly, when study participants are performing tasks that require them to make trust decisions, their decisions are of little value if they do not believe they are legitimately at risk. However, ethical guidelines prevent us from putting study participants in actual danger (in addition to creating subsequent participant recruitment problems). These concerns create another set of lies that we must tell our study participants to yield scientifically valid results.

In this paper, we argue that deception is often required when conducting usability studies of online privacy and security systems. We specifically discuss two types of lies that we tell study participants:

1. Priming can be minimized by deceiving participants about the purpose of the study and introducing subterfuge tasks.

2. Observed trust decisions are only generalizable when participants are led to believe they are at risk.

In the next section, we explain why priming is especially a problem for usable security researchers and why deceiving study participants is necessary to minimize priming effects. We describe several usability studies that our team has conducted that involved deceiving participants to think they were at risk, in order for us to yield valid results. These studies were in the areas of website privacy policies and web browser phishing warnings. While these studies were conducted in our laboratory, the purpose was to gain a better understanding of how users behave online.

Studies

Over the course of the past five years, we have designed and conducted several studies to examine users' online privacy and security perceptions, as well as how they interact with systems designed to enhance their online privacy and security. In this section we provide overviews of our methodologies.

Privacy Premiums

In 2004, we developed Privacy Finder, a new search interface that displays privacy information as search result annotations.[1] This way, web users can make decisions about

[1] http://www.privacyfinder.org/

which website to visit based on privacy policies. To examine the effectiveness of this interface, we conducted a series of usability studies [3, 4, 2].

Due to the aforementioned problems with simply asking people to state their privacy preferences, we did not wish to prime participants to the purpose of these studies. Therefore, we advertised each experiment as an "online shopping and searching study." When participants arrived at our laboratory, we told them that we were generally interested in how they interact with search engines when making online purchases, and we would therefore be observing them use our custom search engine. So as to minimize priming effects, we changed the name from "Privacy Finder" to simply "Finder." We created a cost for increased privacy by preselecting the search results such that purchasing from the high-privacy merchants cost more. Since we paid participants a static amount for their participation, the premium cost of higher privacy came directly out of the participants' pockets.

We created an experimental condition that annotated search results with icons representing privacy levels, as well as a control condition where these icons were absent or relabeled to represent irrelevant information. At no time did the experimenter discuss the icons or privacy itself, though a printed screenshot annotating the search engine features was provided in a packet of materials to each participant for their reference. When using the computer in their natural environments, no experimenter is present to prompt shoppers about the search results that are most in line with their privacy preferences, however, they may have access to help files.

We included subterfuge tasks that involved searching for product information (e.g., "what is the average cost for a pair of Ugg boots?"). This was to both familiarize participants with the search interface, but also to deceive them into thinking that the purchasing tasks were not the only thing we were studying.

Finally, we were concerned that if participants did not believe they were facing legitimate privacy risks, they would not pay any attention to privacy information—why should they? For this reason we required participants to make actual purchases from unfamiliar (but real) merchants. Participants used their personal credit cards and billing information so that their concerns for privacy would approximate the concerns they would have when making purchases under normal circumstances. In this manner, participants understood that the risks they faced in our laboratory were the same as the risks they faced in their natural environments when making online purchases.

Phishing Warnings

Security warnings are a web browser's last line of defense against many of the online threats that face users. These warnings attempt to alert users to potential phishing websites, man-in-the-middle attacks, or other types of insecure websites. When users encounter these warnings, they are often in a mindset to fall for an attack. For instance, when a user views a phishing warning after clicking a link in a fraudulent email, she incorrectly trusts the email and is prepared to transmit her credentials to the phishing website. To properly study these warnings, study participants must be in a similar mindset. We conducted a study to examine the usability of current web browser phishing warnings [1]. This study required particular attention to study design to minimize priming effects and to simulate participants' natural environments.

The first problem we encountered was framing the study so that participants were not primed to phishing concerns. To do this, we advertised the study as another "online shopping study," and told participants that we would be observing their purchasing behaviors. Each participant purchased items from Amazon and eBay using his or her own billing informa-

tion. After each purchasing task was completed, the experimenter provided the participant with a survey on her shopping experience. This survey served as subterfuge while the experimenter sent the participant a phishing message spoofing either Amazon or eBay. Before the participant proceeded to subsequent tasks, she was asked to check her email for the order confirmation—at which point she also encountered the phishing message, subsequently followed the link to the spoofed website, and then encountered a security warning. While this particular attack was highly targeted (participants were more likely to believe the spoofed emails because they had just done business with the website in question), it put participants in the same mindset as they would have been when viewing a phishing warning in their natural environments: they viewed the email as legitimate, and then the web browser warning was the last defense against viewing the phishing website.

The second problem that we needed to address was creating an actual sense of risk, such that participants would be forced to make a value judgment (e.g., is it worth ignoring this warning?). To approximate a real phishing attack, we registered two domain names and designed websites that were indistinguishable from actual phishing websites. This was the closest approximation we could make, since using real phishing websites would have been unethical due to the severe risks that would place on participants. Since participants were under the impression that we were studying the usability of shopping websites, they did not believe that the warnings were part of the experiment. Participants therefore had reason to believe there was a potential risk when ignoring the warnings, and thus we approximated the conditions in which they would be viewing similar warnings in their natural environments.

Conclusion

To best evaluate the effectiveness of online privacy and security systems and interfaces, researchers must attempt to capture how users interact in their natural environments. This is difficult because users often say they are very concerned about their privacy and security, but act in ways that are not consistent with their concerns. We believe that in order to yield valid study results, we must deceive participants as to the purpose of the study and by creating an environment where users perceive that they are subject to real risk.

References

[1] S. Egelman, L. F. Cranor, and J. Hong. You've been warned: An empirical study of the effectiveness of web browser phishing warnings. In *Proceedings of the ACM Computer-Human Interaction Conference*, New York, NY, USA, April 2008. ACM Press.

[2] S. Egelman, J. Tsai, L. Cranor, and A. Acquisti. Timing is Everything? The Effects of Timing and Placement of Online Privacy Indicators. In *Proceedings of the ACM Computer-Human Interaction Conference*, New York, NY, USA, 2009. ACM Press.

[3] J. Gideon, S. Egelman, L. Cranor, and A. Acquisti. Power Strips, Prophylactics, and Privacy, Oh My! In *Proceedings of the 2006 Symposium on Usable Privacy and Security*, pages 133–144, 12-14, July 2006.

[4] J. Tsai, S. Egelman, L. Cranor, and A. Acquisti. The effect of online privacy information on purchasing behavior: An experimental study. In *Proceedings of the 2007 Workshop on the Economics of Information Security (WEIS'07)*, Pittsburgh, PA, USA, 2007.

[5] A. Whitten and J. D. Tygar. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. In *Proceedings of the 8th USENIX Security Symposium*, August 1999.