Trust Me: Design Patterns for Constructing Trustworthy Trust Indicators

Serge Egelman

April 2009
CMU-ISR-09-110

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Lorrie F. Cranor, chair
Jason I. Hong
James D. Herbsleb
Steven M. Bellovin, Columbia University

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

This research was sponsored by the National Science Foundation under grants CCF-0524189 and CNS-0831428, U.S. Army Research Office contract no. DAAD19-02-1-0389 to Carnegie Mellon University's CyLab, and by Microsoft Research.
The views and conclusions contained in this document are those of the author, and should not be interpreted as representing the ofcial policies, either expressed or implied, of any sponsoring institution, the U.S. government, or any other entity.

Keywords: Trust indicators, phishing, security warnings, web browsers, privacy, SSL

Abstract

In a world where making an incorrect online trust decision can mean giving away highly personal information to a con artist, Internet users need effective online trust indicators to help them make better trust decisions. In a perfect world, software could automatically detect all security threats and then block access to high risk websites. Because there are many threats that we cannot detect with 100% accuracy and false positives may exist, web browser vendors choose to warn users about security threats.

Privacy threats also abound on the Internet, but unlike security threats, concerns about privacy threats are nuanced; not everyone cares what a website may do with personal information. To address the varying privacy needs of Internet users, privacy information can be conveyed using contextual indicators that represent privacy policies, because natural language privacy policies are notoriously difficult to read.

In this thesis I qualitatively examine online trust indicators across three varying contexts: web browser phishing warnings, web browser SSL warnings, and indicators that represent website privacy policies. I create guidelines for overcoming many common trust indicator failures, and then I validate these guidelines. I examine these different contexts using a model from the warning sciences in order to shed light on how common failures can be avoided and how design concerns change based on context. I used the results of several user studies that I conducted to compile a set of design patterns for online trust indicators that help designers overcome many common indicator failures. Finally, I highlight the different design considerations between high risk warnings and contextual indicators.

To all the lazy software developers who made this thesis ~~possible~~ necessary.

Acknowledgments

Back in graduate school, I'd learned how to survive without funding, power, or even office space. Grad students are lowest in the academic hierarchy, and so they have to squeeze resources from between the cracks. When you're last on the list for telescope time, you make your observations by hanging around the mountaintop, waiting for a slice of time between other observers. When you need an electronic gizmo in the lab, you borrow it in the evening, use it all night, and return it before anyone notices. I didn't learn much about planetary physics, but weaseling came naturally. –Clifford Stoll, The Cuckoo's Egg

Throughout the years, many people have helped turn me into a better weasel. As far back as I can remember, I was ingrained with the belief that education has the power to cure many of society's ills. I grew up in an environment where intellectual curiosity was highly valued, academic rigor was a virtue, and the pursuit of knowledge was a joyful and never-ending quest. While other children idolized sports icons, I wanted to be a scientist. Unfortunately, in our current world, higher education remains a privilege and not a right, yet I feel very fortunate to be so privileged. For this, their love and support, I am grateful to my parents.

The best advice I received when applying to graduate schools was to look for a good advisor rather than a good school or program. The dynamic between student and advisor has the power to make or ruin a graduate school career. Lorrie Cranor has been a truly excellent advisor and is responsible for many of my graduate school successes. After working with several other collaborators, I have found that her attention to detail has rubbed off on me; I am sure they are just as frustrated with me as I was with her when corrected on grammar or engaged in long drawn-out debate over study methodologies. As much as I joke about my indentured servitude, I am extremely grateful that she set reasonable workload expectations, was available to give advice whenever I needed it, provided me all the resources I needed to be successful, and rewarded me with ample vacation time. Of course, I make these observations from a sample size of only one, for which I also credit her.

I owe a great debt to the other members of the CUPS Laboratory for their feedback over the years. Rob Reeder has been a great friend and peer. He has been a great source for advice throughout graduate school, and specifically the thesis process, as well as a great friend for keeping me sane outside of school. While we joke that I am bitter that he beat me to being Lorrie's first graduated student, I am glad to have had the opportunity to learn from him. I look forward to many future collaborations, academic and otherwise. Janice Tsai has been a great colleague and co-author. Her persistence can be credited for many of our successes, and I look forward to many more future arguments with her over authorship. In alphabetical order, I want to thank Mandy Holbrook for great advice on wording survey questions, Patrick Kelley for his pragmatism and help in keeping things in perspective, Ponnurangam Kumaraguru for his attention to

detail and advice on presentations, Aleecia McDonald for meticulously proof-reading papers, Steve Sheng for his valuable insight and feedback during meetings, and Kami Vaniea for taking over administration of our laboratory server so that I had more time for other endeavors. I also want to thank several faculty members: Alessandro Acquisti for assistance with statistics and his economics perspective, Lujo Bauer for his perspective as a systems person, Julie Downs for guidance with surveys and advice on designing human subjects experiments, Jason Hong for his perspective as an HCI person, and Norman Sadeh for posing intriguing questions during presentations. Finally, I want to thank the members of my committee for their patience, guidance, and thoughtful feedback during the thesis process.

Stuart Schechter at Microsoft Research has been a great colleague, a proponent of my research, and most importantly a friend. I look forward to our continued collaborations. I am thankful to Jeffrey Friedberg for his hard work getting me involved with the Trust User Experience (TUX) Advisory Board at Microsoft, as well as his assistance in reaching out to product groups. I also want to thank Jeb Haber and Jess Holbrook for their assistance and resources. A.J. Brush and Kori Inkpen were instrumental in initially opening the door for me at Microsoft; I want to thank them for a very successful internship. I also want to thank Jim Thornton at PARC for my very first research internship; I learned a great deal about industry research and had a very enjoyable summer.

I thank my sister, Liana Egelman, for providing necessary entertainment during the much-needed vacations we took during graduate school. Finally, I owe a lot to Carolyn Denomme for putting up with me during this process. I hope that I can provide her with the same support and understanding that she has provided me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

During the infancy of the Internet, users generally had technical backgrounds, there were fewer attackers than there are today, as well as fewer software vulnerabilities. This has changed now that 73% of the US population uses the Internet [87]; Internet users are no longer expected to be savvy. Since most Internet users cannot manually detect every conceivable security threat, security software has been developed to protect them. Security practitioners generally recommend that users install a plethora of security software to protect them from today's threats. Anti-virus software protects users from malicious executables that inadvertently get onto their computers due to software vulnerabilities or as a result of making poor trust decisions. Most web browsers now include many different security features: SSL ensures that eavesdroppers cannot decode private data such as credit card numbers or authentication credentials, pop-up blocking features prevent users from being annoyed by unwanted advertisements, and phishing detection features prevent users from being tricked into entering personal information at malicious websites. Privacy features also exist: many web browsers allow users to block cookies, many websites post privacy policies, and several different "privacy seal" programs exist for website owners to convey trust and accountability through the display of graphics indicating their membership.

While a variety of security software and security features exist, they are usually not at the forefront of most users' minds. Privacy and security software is rarely used to facilitate a primary task; users do not sit down at their computers to "do security." Instead, privacy and security software usually runs in the background, only communicating risks via "trust indicators." Trust indicators are displayed when an action was taken on the user's behalf, when the user must make a decision, or to provide additional contextual information. In a perfect world, privacy and security software would act autonomously with no need for user intervention. Because all dangers cannot be detected automatically and mitigated with 100% accuracy, the software must assist users in making trust decisions. However, many users still make poor online trust decisions, which indicates that current trust indicators are failing users. Indicators may fail users for a variety of reasons, some examples include:

- Users may fail to notice the indicators.
- Users may not understand what the indicators represent.

- Users may not trust the indicators.
- Users may be habituated to the indicators.

Most security indicators displayed by web browsers fall into two categories: active warnings that must get the user's attention so that she is alerted to an impending danger, and passive indicators that display contextual information to aid the user in making informed decisions. In this thesis I examine both types of indicators in order to create guidelines for preventing common indicator failures. Specifically, I examine active web browser indicators (i.e. warnings) that alert users to phishing websites and SSL errors, as well as passive indicators (i.e. icons) that represent website privacy policies.

1.2 Research Questions

> *In this thesis I created and validated a series of design patterns to address many of the common failures of both critical warning messages as well as contextual indicators found in web browsers.*

I present the results of five studies that I conducted in order to examine web browser privacy and security indicators. I conducted these studies to answer the following research questions:

1. How can we design privacy/security indicators used in web browsers to help users make the best[1] trust decisions given available information? *Chapter 9*
2. What are the differences in design concerns for passive and active indicators? *Chapter 9*
3. Does the use of different "severity levels" minimize habituation for the most serious warnings? *Chapters 3, 4, and 5*
4. Does the design of a website cause the user to trust an indicator less? Does indicator effectiveness improve when it is not displayed along with the website content? *Chapters 3 and 6*
5. Are users more likely to make a better trust decision when the indicator highlights the recommended choice (or makes the recommended choice the default or simpler action)? *Chapters 3, 4, and 7*
6. Are contextual indicators more likely to be taken into account when shown before a user chooses to visit a website (i.e. compared to after they choose a website)? *Chapter 6*
7. Does the choice of text in a warning recommendation impact a user's decision to obey the warning? *Chapters 4 and 5*
8. Are users more likely to take a safer action when the warning makes a recommendation (i.e. when compared to a warning that does not recommend a course of action)? *Chapters 3 and 4*

[1] Best is defined as maximizing the user's utility function.

1.3 Overview of Studies

I conducted two studies on phishing warnings and one study on SSL warnings. All of these warnings were "active" warnings—they forced the user to react and choose between several options in order to continue. I also conducted two studies on passive indicators which provided contextual information regarding website privacy policies. Based on the results of these studies, I constructed a set of design patterns to aid developers in creating future trust indicators.

1.3.1 Phishing Warning Study

The newest web browsers now include active phishing warnings, which force users to notice the warnings by interrupting their tasks. In order to examine whether users are likely to obey these warnings Internet Explorer 7 (IE7) and Firefox 2 (FF2), I performed a laboratory experiment where I observed participants' reactions to both active and passive phishing warnings. I simulated a "spear phishing" attack by sending participants phishing messages after they had just completed online purchases. I examined the differences between active and passive warnings by triggering IE7's two different phishing warnings. Participants in one condition were shown a dialog box that was easy to dismiss and did not offer any options, which I considered to be a passive warning because it provided no options and did not force the user to interrupt her current task. Participants in another condition were shown a full-screen message that blocked the destination website and offered recommended options for how to proceed, which I considered to be an active warning. I also examined the active phishing warning used by FF2. Overall, I found that the active warning used by FF2 was significantly more effective than the active warning used by IE7, because the IE7 users confused the phishing warnings with warnings they had seen in much less risky situations. They therefore became habituated to the warning because it looked similar. I also found that both active warnings were significantly more effective than the passive warning, and the passive warning was not statistically different than the absence of any warning. This indicates that phishing warnings should be prominent, designed differently from less serious warnings, and force the user to interact with the warning.

1.3.2 Warning Options Study

Based on the results of my first study, I found that the IE7 phishing warning was ineffective because it looked like other less-serious IE warnings. I performed a second study on phishing warnings to examine whether a new design would minimize this habituation effect, as well as to examine the role of the option text. In each condition I kept the ill-advised option of "disregard and continue (not recommended)," but created two experimental conditions to observe which recommended option resulted in the highest rate of compliance. In one condition participants were advised to "go to my homepage," while in the other condition participants were advised to "find the correct website." The purpose of the two varying text options was to examine whether an option that appeared more conducive to finishing the task—"find the correct website"—would be more attractive to study participants. To distinguish this warning from other warnings in IE, I tested a red border and used a version with a white border as a control condition. I examined whether participants would pay more attention to a warning that appeared more severe—indicated by a red border—and consequently make safer choices. Overall, I observed an interaction effect between the red background and the option text: participants who saw warnings with red borders and the option to "find the correct website" spent significantly more time viewing the warnings. However, ultimately I observed no significant differences in

compliance with the warning (i.e. not proceeding to the destination website) because the rest of the warning text did not adequately convey the risks and consequences of proceeding.

1.3.3 SSL Warning Study

I conducted a third study on active warnings to validate my previous findings by using the previous findings to create and test a new warning for alerting users to SSL errors. In this study I examined current SSL warnings for self-signed certificates alongside this new warning that I developed. Previous SSL warnings often fail for one of two reasons: they do not adequately convey risk, resulting in users overriding them in high-risk situations; or they appear very dire or make it difficult for users to override them, which results in users being unable to visit legitimate websites in the event of a false positive. I designed a new SSL warning to clearly convey risk and consequences, and to use user input to gauge the risk level of a given situation in order to minimize false positives, thereby minimizing habituation. I discovered that when using my new warning, participants were significantly more likely to obey the warning when they were in danger, while also knowing to disregard the warning when it appeared on a legitimate website (i.e. a false positive). Based on my exit survey, I found that this was because significantly more participants who saw my custom warning understood the risks, as compared to those who saw the warnings used by FF2, FF3, and IE7. While this study showed how to improve warnings, it also showed that warnings are an imperfect solution to security problems: even when exposed to improved warnings, half the participants still made poor choices.

1.3.4 Privacy Information Timing Study

I conducted a study on passive privacy indicators to examine how the timing of their appearance impacts user behaviors. Additionally, I examined whether users pay less attention to passive indicators when they are displayed alongside website content. In this study, participants made online purchases using a search engine interface that provided privacy information about the resulting websites. Participants saw the privacy indicators displayed in one of four ways: irrelevant information substituted for privacy indicators (the control condition), as search result annotations, as a frame above the destination website after clicking a search result, or as an interstitial seen after clicking a search result but before seeing destination website. Overall, participants who saw privacy indicators, regardless of when they were displayed, paid significantly more money for higher privacy when they were purchasing privacy-sensitive items. Additionally, by annotating search results, participants were able to locate the high-privacy websites significantly quicker than those who saw the privacy indicators only after selecting a website from the list of search results.

1.3.5 Privacy Finder Usage Study

I further examined passive indicators that represented website privacy policies using Privacy Finder, a privacy-enhanced search engine that I designed. In the previous study on passive indicators, I found that search result annotations were an effective way of communicating privacy information. As such, I created Privacy Finder to annotate websites with passive indicators rating the strength of a given search result's privacy policy. I performed a field study using Privacy Finder to examine whether participants would take passive privacy indicators into account when choosing websites to visit. I tested this by examining whether participants visited search results further down the list in order to find websites with better privacy policies. Given a random set of search results, if the indicators did not alter browsing behaviors then the distribution

of search result visits should be similar for those annotated with privacy indicators to those not annotated with privacy indicators. Overall, I found that this was the case: participants were significantly more likely to click search results when they were annotated with privacy indicators. I also observed that search results that appeared further down the page (i.e. beyond the first result) were significantly more likely to be visited when they were annotated with privacy indicators.

In Chapter 2 I present related work on trust indicators. In Chapter 3 I present the results of the Phishing Warning Study. In Chapter 4 I present the results of the Warning Options Study. In Chapter 5 I present the results of the SSL Warning Study. In Chapter 7 I present the results of the Privacy Finder Usage Study. In Chapter 6 I present the results of the Privacy Information Timing Study. Finally, in Chapter 8 I present the design patterns that I created based on the results of the four previous studies. I conclude with Chapter 9.

Chapter 2

Background and Related Work

Security and privacy problems have existed on the Internet since its inception, but have gotten much worse in recent years, largely due to the growth of the Internet. These threats have major impacts on end users, websites, businesses, and even Internet service providers. Technical solutions exist for many problems, however, users are still forced to make online trust decisions. The consequences for making a poor online trust decision can result in financial loss, identity theft, or even destruction of property.

Trust indicators exist to communicate security and privacy information to users so that they can make informed decisions. Trust indicators can alert users to actions taken on their behalf, advise them on recommended actions, or simply provide contextual information. In this chapter I discuss several online security and privacy issues, some proposed solutions involving trust indicators, some usability studies of these proposed solutions, and then describe a model from the warning sciences that I apply to improve both future and existing solutions.

2.1 Online Security Threats

Researchers estimate that the time it takes between a Windows machine being plugged into a network and it being compromised is a matter of minutes [125]. While this problem may be due to flaws in the operating system or other software, many online security problems exist because users are not given the best tools to adequately understand risk [69]. Another aspect of this problem is that users often have different mental models of how computers operate [79]. Users have an especially hard time trying to grasp security concepts.

To give an example, in a 1999 user study of the PGP 5.0 email encryption system, researchers found that the usability of this security software was woefully inadequate. One of the twelve participants in the study was never able to figure out how to encrypt email, while those who did figure out the encryption process took up to thirty minutes. Even then, only two participants were able to use the correct keys for encryption. The authors concluded that "it is clear that there is a need to communicate an accurate conceptual model of the security to the user as quickly as possible" [137].

Despite the threats to users and the fact that many claim to value security, users are often willing to give up security in exchange for other benefits. In one study of a piece of software called "Polaris," researchers found that a majority of participants were willing to ignore security precautions in order to quickly complete the primary task [39]. In another study that validated what many usability researchers have believed to be

true for many years, researchers found that when confronted with dialog boxes, most users will dismiss them without reading in order to continue their primary tasks [95].

2.1.1 Semantic Attacks

Current computer security software cannot automatically detect all threats and automatically act on the users' behalf. As a result, humans are required to make security decisions. Attackers can and will exploit these decisions because it is often easier than exploiting a technical vulnerability. One class of threats, semantic attacks, rely on tricking users into divulging personal information [111]. Phishing is an example of a semantic attack. The Anti-Phishing Working Group (APWG) reports that as of January 2007, 29,930 unique phishing URLs were reported to them [10]. Phishing is partially responsible for the dramatic rise in identity theft, which cost consumers over half a billion dollars in 2004, according to the U.S. Federal Trade Commission (FTC) [51]. The cost to banks and card issuers from phishing attacks is in the billions of dollars [86]. The susceptability of consumers to fall for phishing attacks comes as little surprise; an informal 2004 study found that 70% of participants would be willing to divulge their passwords for a bar of chocolate. In fact, 34% of those surveyed divulged passwords before being offered any incentives [13]. Of course, there is no way of knowing whether or not participants divulged their real passwords, but this example supports other research highlighting people's willingness to exchange private information for upfront incentives. This problem may be attributed to a misunderstanding of risk, or an altered perception of risk due to the upfront incentive [3, 4].

Despite growing efforts to educate users and create better detection tools, users are still very susceptible to phishing attacks. Unfortunately, due to the nature of the attacks, it is very difficult to estimate the number of people who actually fall victim. Victims may not disclose how much they lost, whether they were victims, or they may not even be aware that they were victimized. Despite this, there have been various attempts to quantify the cost of phishing. A 2006 report by Gartner estimated the costs at $1,244 per victim, an increase over the $257 they cited in a 2004 report [63]. In 2007 Moore and Clayton estimated the number of phishing victims by examining web server logs. They estimated that 311,449 people fall for phishing scams annually, costing around 350 million dollars [92]. Another study in 2007 by Florencio and Herley estimated that roughly 0.4% of the population falls for phishing attacks annually [54].

Phishing works because users are willing to trust websites that appear to be designed well. In a 2001 study on website credibility, Fogg et al. found that the "look and feel" of a website is often most important for gaining a user's trust [55]. A 2006 phishing study by Dhamija et al. found that 90% of the participants were fooled by phishing websites. The researchers concluded that current security indicators (i.e. the lock icon, status bar, and address bar) are ineffective because 23% of the participants failed to notice them or because they did not understand what they meant [42]. In a similar study, Downs et al. showed participants eight emails, three of which were phishing. They found that the number of participants who expressed suspicion varied for each email; 47% expressed suspicion over a phishing message from Amazon, whereas 74% expressed suspicion over a phishing message from Citibank. Those who had interacted with certain companies in the past were significantly more likely to fall for phishing messages claiming to be from these companies. Participants were also likely to ignore or misunderstand web browser security cues [44].

Dhamija et al. conducted a study in 2006 that examined how users examined phishing websites. They found that 23% of the participants did not look at any of the browser's security features, resulting in them being tricked 40% of the time. A total of 90% of the participants were fooled during the course of the

study [42]. In a similar study, Downs et al. showed participants eight emails, three of which were phishing emails. They found that the number of participants who expressed suspicion varied for each email; 47% expressed suspicion over a phishing message from Amazon, whereas 74% expressed suspicion over a phishing message from Citibank. Participants who had interacted with certain companies in the past were significantly more likely to fall for phishing messages claiming to be from those companies. Participants were also likely to ignore or misunderstand web browser security cues [44]. Thus it is likely that traditional browser security indicators are inadequate, and that improved trust indicators are needed. Developers of web browser software could aid users by creating effective trust indicators that alert users to potential phishing attacks.

2.2 Online Privacy Concerns

Privacy is often cited as a top concern among Internet users [2]. According to a 2005 poll conducted by CBS News and the New York Times, 82% of Americans believe that the right to privacy in the U.S. is either under serious threat or is already lost. This same poll also found that 83% of Americans are concerned about companies collecting their personal information because of the risk that companies might share their personal information inappropriately [22]. A 2008 poll conducted by Consumer Reports indicates that 72% of Americans are "concerned that their online behaviors were being tracked and profiled by companies" [28]. These responses are similar to a 2000 survey conducted by The Pew Internet & American Life Project, in which 86% of respondents said that they wanted companies to require permission before using personal information for purposes other than those for which it was provided [56].

In response to consumer privacy concerns, many corporations have posted privacy policies [73]. But these policies rarely help because they often go unread [91], or do not address the most common consumer concerns [45, 103]. Furthermore, a majority of individuals surveyed held the mistaken belief that the mere presence of a privacy policy means that a corporation will not share their data [130]. Even those who do bother to read privacy policies often cannot understand what the policies mean [35]. Anton et al. examined forty bank privacy policies and found that on average, a college education was needed to comprehend them [11]. A 2008 survey found that several years of graduate school are required to read the privacy policies of the top Internet companies [113]. In another 2008 study of the privacy policies found on 75 popular websites, researchers showed that it would take an average of ten minutes to skim each one of them. This study estimated that Internet users who read the privacy policies at web sites they visit just once per year would spend over 200 hours per year reading privacy policies [89].

2.3 The Usability of Trust Indicators

There is evidence to show that users need tools to help protect them from online privacy and security threats. At the same time, these tools need to be designed such that they convey the most relevant information in an intuitive manner in order for users to make informed trust decisions. In this section I provide some examples of trust indicators used to protect online users from current security and privacy threats. I specifically survey indicators used for SSL, authentication, phishing website identification, and website privacy policies.

2.3.1 Security Indicators

SSL

One of the most common trust indicators found on the World Wide Web is the SSL key and/or lock icon. These icons appear within a web browser to inform users that their connections are being encrypted with SSL or TLS [106]. However, these indicators traditionally say nothing about the legitimacy of a site, only that the connection is encrypted. Thus, even fraudulent websites can and do use legitimate SSL certificates [52, 100]. These indicators are designed to be displayed within the chrome of the browser, such that they cannot be altered by website content. However, various software vulnerabilities have allowed malicious code to modify browser chrome [85, 6, 150]. But, a successful attack does not need to modify the browser chrome. A study in 2002 found that half of the participants could not identify a secure browser connection [58]. This number is likely an upper bound for noticing security indicators because participants were primed for security—they were specifically asked to identify the security indicators. Another study in 2005 used eye trackers and found that when using a web browser, participants paid no attention to security cues (such as SSL icons) within the web browser. Only after priming participants to be on the lookout for security information, 69% of them noticed the lock icons [135]. Thus it is unlikely that most Internet users in their natural environments notice current SSL indicators when using web browsers. I build on this research in this thesis by examining SSL warning messages: if users cannot be expected to notice the presence of SSL icons, instead we should focus on creating noticeable SSL warnings in the event that they encounter a problem.

Authentication

In addition to using SSL to identify an encrypted website, SSL can also be used for mutual authentication. Mutual authentication involves both the server identifying itself to the client and the client identifying itself to the server. For example, when the client shares a certificate signed by a trusted third party, the server can trust the identity of the client. Mutual authentication is one possible solution to the phishing problem. However, it is not clear that SSL is the best solution. In 2004, Marchesini et al. identified a myriad of methods for compromising a user's private key [74]. But assuming that the technical aspects can be fixed, using SSL for mutual authentication creates many usability problems, largely because users have difficulty understanding how public key cryptography works [137].

There are many problems with standard password authentication mechanisms. The two most common ones are having to remember many passwords for many different sites [117, 5, 148] and not having an accurate mental model for how the system works [137, 79]. Thus, alternative mechanisms have been proposed. The use of graphics for authentication has been proposed as a way of decreasing the burden of having to memorize a textual password. Several proposals for graphical passwords have been centered around the user drawing a picture to authenticate [78]. Others have focused on users clicking areas within a picture or choosing between different images [25, 16, 123]. Another proposal used randomly generated art that the computer assigned to the user as a password [40]. These systems may be advantageous since the passwords may be less likely to be forgotten and also may be less susceptible to eavesdropping. However, studies have also shown that users take more time to enter graphical passwords [138, 99, 124]. The increase in time it takes to authenticate may actually make these systems more susceptible to eavesdropping. However, more importantly, the increase in time it takes for a user to authenticate is likely to frustrate the user. Since this prevents users from quickly initiating their primary tasks, they may adapt by engaging in unsafe behaviors

such as sharing accounts and logging out less frequently.

The Passfaces system was another graphical authentication system invented to combat some of the usability and security problems of textual passwords by providing the user with human faces to choose from as a means for authentication [29]. One preliminary study found that while participants made fewer errors than with traditional passwords, they spent more time trying to log in, and logged in less often to save themselves the hassle [18]. Another study found that Passface users tended to make predictable choices, largely based on gender and ethnicity [36]. Thus, credentials used by authentication systems need to be hard to guess, yet efficient to use.

The Passpet system, created by Yee et al. in 2006, uses visual indicators for mutual authentication. The tool is a browser extension that uses a trusted path that stores icons of animals within the web browser. Users can store an icon for the trusted sites with which they interact, so that the system will only send a password when the animal icon is the same as the previously chosen one. Preliminary user testing suggests that this system is easy for users to use [151]. However, it does require third party software to be installed, and it is not clear how users would access their online accounts when using different computers without having to install this software. Similar "usable" solutions for improving online security suffer from similar problems. Parno et al. created another mutual authentication system in 2006 that relies on out of band communications. The user uses a cellular phone to securely store certificates, which are exchanged with the user's computer via Bluetooth [101]. While this system is also effective at achieving mutual authentication, it is not clear how usable the security indicators are on the phone and if the user would fall victim to spoofing attacks. Additionally, if the user has misplaced the phone or if the software is not installed on the computer, authentication cannot take place.

The SiteKey system was invented in 2005 to solve some of the problems with mutual authentication. SiteKey uses a system of visual mutual authentication images that are selected by the user at the time of enrollment. When the user returns to a website, a username is entered, at which point the stored image is displayed. If the user recognizes the image as the original shared secret, it is safe to enter the password, as it is likely this site is the legitimate one [12]. However, Schechter et al. found that 92% of participants still logged in to the website using their own credentials when the correct image was not present [110]. However, this sample may have been drawn from a biased population since others refused to participate, citing privacy and security concerns. At the same time, it is clear that relying on users to notice the absence of a security indicator is destined for failure. Therefore, in this thesis I show how passive indicators should only be used to convey contextual information; when users are confronted with an impending danger, an active warning should be used instead.

Phishing

Phishing is a specific type of semantic attack and is another area where trust indicators can be used to help users identify trusted websites. One method for identifying trusted websites which is currently being aggressively marketed by certificate authorities is the use of *extended validation* (EV) certificates. An EV certificate differs from a standard SSL certificate in that the corporation purchasing the certificate must go through more rigorous background checks. A regular certificate only tells a user that the certificate was granted by a particular issuing authority, whereas an EV certificate also says that it belongs to a legally recognized company [23]. Most of the major web browsers now include special support for EV certificates. For instance, the newest version of Microsoft's Internet Explorer will color the URL bar green and display

the name of the company (Figure 2.1). Similarly, the TrustBar extension, available for Firefox and Mozilla, also displays certificate information within the browser chrome [94]. However, a recent study found that EV certificates did not make users less likely to fall for phishing attacks. Many users were confused when the chrome of the web browser was spoofed within the content window to depict a green address bar. Additionally, the study found that after reading a help file, users were less suspicious of fraudulent websites that did not yield warning indicators [76]. In 2008, Sobey et al. performed a study on EV indicators using eye trackers. They found that none of their participants noticed—much less interacted with—the EV indicators when performing online shopping tasks [116]. Yet similar indicators are still being used by both Internet Explorer and Firefox, despite their proven ineffectiveness.

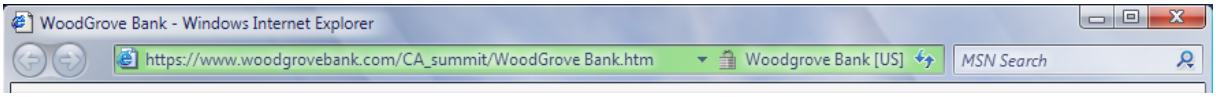


Figure 2.1: Screenshot of the IE7 web browser depicting an EV certificate [57].

Passive indicators have been displayed by tools that notify users of potential phishing websites. Unfortunately, in a study I performed in 2006 on anti-phishing tool accuracy, we found that these tools failed to identify a substantial proportion of phishing websites [154]. The usability of these tools is also lacking. In 2006, Wu et al. tested the effectiveness of the indicators used by some of the more popular anti-phishing tools. They found that many users failed to notice the indicators. Many of those who did notice the indicators did not trust them because they believed the tool was in error since the website looked trustworthy [146]. The factors that go into website trust have been extensively studied by Fogg et al., who found that the "look and feel" of a website is often most important for gaining a user's trust [55]. Thus it is easy to understand why a user might trust a professional looking website despite the presence of a passive warning indicator that is displayed in the browser chrome.

Other proposals have been put forth to modify browser chrome to help users detect phishing websites. In one system, *synchronized random dynamic boundaries* (SRD), by Ye and Smith, the browser chrome is modified to blink at a random rate. If the blink rate matches a trusted window's blink rate, the user knows that the window in question has not been modified by a malicious website [149]. While effective during preliminary user studies, this system requires extensive browser modifications, and thus may not be a feasible solution. Furthermore, this system was not tested under realistic conditions where a user may already have several windows open that may distract from the blinking border. A similar solution—with similar caveats—that uses a trusted window that must be compared with the browser chrome was also proposed by Dhamija and Tygar in 2005. In their system the chrome of the browser window contains a colored pattern that must be matched with the trusted window. The user knows to recognize the trusted window because it contains a personal image that the user selected during the initial configuration [41]. All of these proposals require users to install third-party tools, thus they have not seen widespread adoption. But it is not clear that they would be effective in the real world: in many of these user studies, the participants were primed for security, they were aware that they were testing a security tool, and they may have felt compelled to behave a certain way. Thus, these results may have been tainted by either the Hawthorne or Milgram effects. The Hawthorne Effect occurs when study participants understand the objective of the study and therefore alter their behaviors accordingly. The Milgram Effect occurs when study participants perform actions against their better judgment because they were instructed to do so by an authority figure [82, 90].

In this thesis I show how these effects can be minimized through proper attention to study design.

2.3.2 Privacy Indicators

Just as with security information, indicators can be used to distill privacy policy information into intuitive icons. However, studies have shown that privacy and security indicators can fail users when they go unnoticed, when they force users to take extra steps to complete a task, or when other environmental stimuli outweigh the strength of the indicators [146, 55, 48]. Previous studies have shown that users may be willing to pay a premium to know when they are visiting a high privacy website [129]. But there is still an open question of *how* to effectively convey website privacy information. I explore the role of privacy indicator design further in Chapter 6.

Privacy Seals

Because of the problems with natural language privacy policies, companies have started to take proactive steps to make themselves seem more privacy-conscientious. Many companies post "privacy seals" on their websites in an attempt to improve consumer confidence. In 2001, Adkinson et al. estimated privacy seal adoption at 11% [7], while Jensen et al. estimated privacy seal adoption at around 2% in 2006 [77]. For FY2008, TRUSTe claimed 3,440 participating websites worldwide, including 24 Fortune 500 participants [127]. Assuming privacy seals are pervasive enough to be recognized by consumers, do consumers properly understand what they represent?

Many Internet users erroneously believe that websites must adopt consumer-friendly privacy practices in order to post these seals. However, the presence of a privacy seal says nothing about the content of a company's privacy policy [93]. In fact, Edelman conducted a study of websites brandishing the TRUSTe privacy seal in 2006 and concluded that "sites that seek and obtain trust certifications are actually significantly less trustworthy than those that forego certification [46]."

If trustworthy privacy seals do exist, it is unlikely that users recognize them. In a study conducted in 2005, 15% of participants claimed to recognize an authentic-looking privacy seal created solely for the purpose of the study. At the same time, the legitimate privacy seals were only recognized by 26% of the participants on average [93].

P3P

The W3C's Platform for Privacy Preferences (P3P) was created to help users understand website privacy policies. P3P specifies a standard set of XML elements that can be used to construct machine-readable privacy policies. These policies can be posted on websites and then analyzed by user agents on behalf of Internet users. If a user agent encounters a privacy policy that does not conform to a user's stated privacy preferences, the user agent can take actions on behalf of the user such as displaying a warning, rejecting cookies, or blocking the website entirely [30]. Byers et al. found that by 2003, P3P had already been adopted by over 30% of the most popular websites and 10% of their entire sample [20]. By 2005, Egelman et al. reexamined this sample and found that P3P adoption had increased by over 30%. They also found that on average, 32% of all Google queries yield at least one P3P-enabled search result [47]. In 2006, Jensen et al. compiled a sample of over 26,000 websites from around the world and used it to estimate P3P adoption at 25% [77]. The increasing rate of P3P adoption is beneficial to consumers because it facilitates the automatic

dissemination of website privacy information; tools can be developed to distill privacy policies into simple indicators automatically.

In 1999, AT&T began developing their Privacy Bird P3P user agent for Internet Explorer. Privacy Bird displays a colored bird icon in the corner of the web browser to indicate whether a policy matches the user's stated privacy preferences. A red bird indicates a conflict with the user's preferences, while a green bird indicates a compliant policy. Cranor et al. conducted a survey of 309 Privacy Bird users and found that a common complaint was that privacy information was not displayed on many websites. They concluded Privacy Bird was still useful since 88% of the respondents said that being aware of website privacy policies caused them to alter their behaviors. Many claimed that they stopped visiting certain websites, sought opt-out information, and compared websites based on privacy policies [34]. However, a shortcoming of Privacy Bird is that to view a website's privacy information, users must first transmit certain clickstream data to visit that website. This also means that to compare the privacy policies of n different websites, a user must visit all n websites before making a decision. It is unclear whether or not a user will go through this process until he or she finds a satisfactory privacy policy.

Privacy Finder

The idea of a P3P-enabled search engine was proposed by Cranor et al. in 2004. Their prototype allowed users to enter a set of search terms and retrieve a list of results annotated with red or green birds indicating whether or not each result complies with the user's stated privacy preferences [33]. Egelman et al. improved this search engine, named it Privacy Finder, and made it publicly available. One of the improvements was the addition of "privacy reports." Users of Privacy Finder can click on the privacy indicators to generate a summarized version of a website's privacy policy highlighting any conflicts it may have with the user's privacy preferences [47].[1] Gideon et al. conducted a user study of Privacy Finder in 2006. Participants were instructed to purchase a privacy-sensitive item—condoms—and a common household item—a power strip. The search results for each product were pre-selected so that at least one green bird icon appeared along with several red bird icons. The websites were selected such that those with better privacy policies had higher prices. Thus, users had to pay a premium for higher privacy. The researchers found that when purchasing the privacy-sensitive item, participants paid significantly more for it than those in a control group who did not see the privacy indicators [66].

We performed a followup to Gideon et al.'s study in 2007. To determine whether participants cared about privacy or were visiting websites simply because they liked the indicators but did not know what they represented, we added a second control condition that used the same indicators as in the experimental condition, but labeled them as representing irrelevant information rather than privacy. We also changed the privacy indicators from red and green birds to a set of four boxes: the number of boxes colored green was inversely proportional to the number of conflicts with the user's privacy preferences; four colored boxes indicated a privacy policy completely matched a user's privacy preferences. We removed the indicator from the website with the lowest price to test the effect of encountering an unknown privacy rating. We conducted an online survey to identify products that would raise participants' privacy sensitivities, but would unlikely result in participants dropping out of the study if asked to purchase them. We chose a vibrating sex toy as the privacy-sensitive item and a pack of AA batteries as an item that would be unlikely to raise privacy concerns [129].

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

We observed that participants were willing to pay a premium to buy from a website with a privacy indicator, however we did not control the exact amount of the premium or keep it constant between the two products. Thus, it is unclear whether participants would have paid the same premium for the two products. We did not test whether participants would pay a privacy premium when the cheapest website had the worst privacy policy (rather than no privacy indicator). Finally, we never examined how alternate methods of displaying privacy indicators impacted purchasing decisions, we only examined annotated search results. Several of the other studies we have cited show how (not) to display indicators in browser chrome [146, 48, 135, 116], but few studies have offered methods for displaying privacy indicators alongside website content.

2.4 Studies in The Warning Sciences

Computer scientists can stand to benefit from studies in the warning sciences—a subfield of ergonomics—when designing online warnings and indicators. Many studies have examined "arousal strength" and "hazard matching." Arousal strength is defined as the user's perceived risk of ignoring a given warning, whereas hazard matching is the process of ensuring that perceived risk matches actual risk [71]. These factors should be taken into account by software developers when they are designing security features. By using research to guide designs, security features can be incorporated into products from the ground up, rather than adding them in as an afterthought. This practice is likely to increase the overall security of a product [1]. Considering user models during the design phase centers the application around the user, such that usability is considered during all phases of the design process [155, 114]. Thus, applying relevant literature to the design of security warnings is likely to result in better user trust decisions.

In one study, Wogalter and Silver experimented with the arousal strength of various warning words, finding a spectrum of responses to the words. They concluded that different words by themselves elicit responses ranging in severity [143]. A followup study was conducted in 1998 by Wogalter et al. where they examined combinations of warning words and icons. In this study they found that participants interpreted the warnings to represent very different hazard levels than what the designers of the warnings intended [141]. These findings suggest that designers of software warnings need to pay particular attention to the words and icons that they use so that their warnings match the actual hazards. If the warnings are designed in an ad-hoc fashion and convey a lower level of risk, users are likely to endanger themselves by ignoring the warnings.

In one study of current warning messages used in Microsoft Windows, researchers found that using different combinations of icons and text greatly impacted the participants' risk perceptions. Amer and Maris conducted a study to determine how users perceive software hazards based on warning messages and icons. Participants were shown a series of dialog boxes with differing text and icons, and were instructed to estimate the severity of the warning using a 10-point Likert scale. The choice in both icon and warning words greatly impacted how each participant ranked the severity. The researchers also examined the extent to which individuals will continue to pay attention to a warning after seeing it multiple times. The researchers found that users dismissed the warnings without reading them after they had been displayed multiple times. This behavior continued even when using a similar but different warning in a different situation. The only way of recapturing the user's attention was to increase the arousal strength of the warning [9]. This effect is known as habituation. Habituation is the decreased response to a stimulus that a person experiences after repeated exposures [64].

When a person first processes a new stimulus that was unexpected or unique, short-term memory is queried to see if the stimulus can be recognized. After that fails, long-term memory is queried. If the stim-

ulus was indeed novel, it becomes encoded into long-term memory [83]. Thus, subsequent similar stimuli will be recalled from long-term memory creating a diminished response with each exposure. Diao and Sundar examined habituation effects with regard to banner advertisements and found that participants quickly became habituated to even animated advertisements after "exposure to the first couple of web pages" [43]. Thus, Spiekermann and Romanow surmised that it is possible for a person to become habituated to a stimulus after only a single exposure [118].

Another possible explanation for habituation is that if there is no obvious consequence to ignoring a stimulus, there is no need to continue responding to it. This is a learned behavior that occurs over time [65]. This obviously applies to online security warnings: if a user ignores a security warning and has her information stolen, it may take weeks before she notices fraudulent transactions on her credit card statement. More importantly, she is unlikely to associate ignoring the security warning with the fraudulent charges. Thus, ignoring security warnings becomes a learned behavior because there are no immediate consequences, and therefore users quickly become habituated to the warnings.

However, there are still ways of preventing habituation. When a person has become habituated to a warning, his attention can be recaptured if the message is varied, so that it is not confused with the original message to which the person has become habituated [64]. These conclusions were similar to Wogalter and Vigilante's recommendations for minimizing habituation: make the warning more conspicuous, modify the warning, or present the warning only when absolutely necessary [144]. Wogalter and Leonard point out that "habituation indicates that there is some information about the warning in memory" [142]. Thus, a person might see a warning, confuse it with a different warning, and therefore fail to comprehend the new warning simply because it had a similar design.

Researchers in this area have also crafted recommendations to help users better comprehend warnings; increased comprehension is likely to result in increased compliance. Using both pictures and text in warnings increases comprehension of the warning message [38], as well as memory retention [153]. Thus, most new warnings are designed with these considerations in mind. At the same time, new types of warnings are being designed to improve risk communication. However, there is little evidence to suggest that this research is being applied within the realm of computer science [37, 9].

2.4.1 The C-HIP Model

Wogalter proposed the Communication-Human Information Processing Model (C-HIP) for structuring warnings research, as shown in Figure 2.2. He suggests that C-HIP be used to identify reasons that a particular warning is ineffective [140]. The C-HIP model begins with a source delivering a warning through a channel to a receiver, who receives it along with other environmental stimuli that may distract from the message. The receiver goes through five information processing steps, which ultimately determine whether the warning results in any change in behavior.

Cranor proposed a series of questions that should be asked when evaluating security indicators [31]:

Do users notice the indicator? This question addresses the Attention Switch and Attention Maintenance steps. Users focused on their primary tasks "may not notice an indicator that is too small, surrounded by more interesting icons, covered up by other windows, or positioned somewhere on the screen where users seldom look" [31]. An effective indicator must first attract a user's attention and then hold it long enough for the user to comprehend its meaning. To test this, user studies should test whether users notice indicators

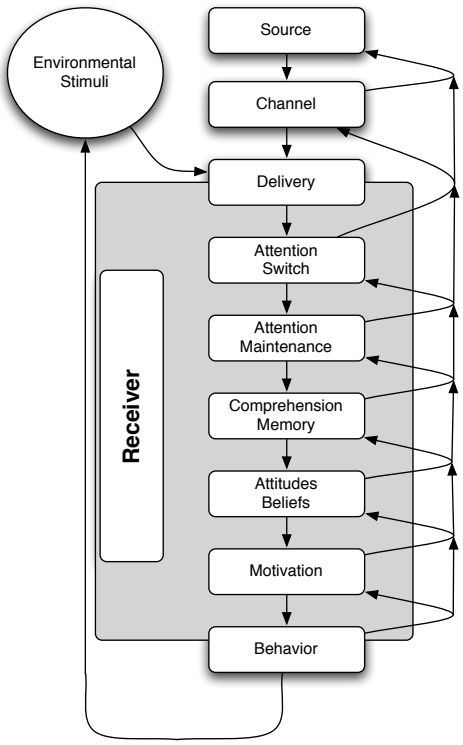


Figure 2.2: Diagram of the different phases of the C-HIP model [140].

while performing other tasks.

Do users know what the indicator means? This question addresses the Comprehension/Memory step in which users understand the indicators meaning for the first time or remember the indicators meaning from a previous encounter. Indicators represented as symbols without text may not be readily comprehendible to users who have not seen them before. Text explanations containing technical jargon might not be comprehensible to non-expert users and text with big words or long sentences might not be comprehensible to users with poor reading skills. Some users might think they know what an indicator means but may actually be misinterpreting the indicator. Thus it is important to evaluate whether target users correctly understand what an indicator means.

Do users know what they are supposed to do when they see the indicator? This question also applies to the Comprehension/Memory step. A user may understand what a particular indicator means, however, they may not understand how they are supposed to react to it. Some indicators may need to suggest that the user take specific actions.

Do they believe the indicator? This question addresses Attitudes and Beliefs about the indicator. Users may understand what they are supposed to do but decide not to do it because they believe that the indicator is unreliable [146].

Are they motivated to take the recommended action? This question addresses the Motivation step. Users may believe that the indicator is reliable, but may nonetheless be unmotivated to actually take the recom-

mended action. Lack of motivation may be due to their perceptions about the risk or its consequences, or about the difficulty or inconvenience associated with the recommended action. Additionally, users' motivations may be altered based on their perceptions about the risks and benefits of following the recommended action.

Do they actually do it? Computer users do not always follow the course of action recommended by indicators, even if they trust the indicators and are motivated. Although empirical warnings studies have shown that people who say they intend to do something often, in fact, do it [115], there are still many cases when people do not do things that they say they intend to do. Don Norman coined the terms *Gulf of Execution* and *Gulf of Evaluation* to describe problems users may have when they attempt to complete the correct action—in this case an action recommended by a warning—but are either unable to or do not receive proper feedback to indicate that the action was completed successfully [96]. These are both failures that occur in the *behavior* stage of the C-HIP model.

Do they keep doing it? This question addresses the Behavior step. After being exposed to a particular indicator over time, users may stop paying attention to it (known as habituation—the loss of ability to facilitate attention switch). It is often very hard to determine long term effectiveness of indicators because the user needs to be repeatedly observed in his or her natural environment. However, this is a very important step to examine because many indicators are rendered useless because they have become the subject of habituation.

How does this indicator interact with other indicators that may be installed on a users computer? The environmental stimuli that influence indicator perception include other indicators on a users computer. An effective trust indicator on its own may be rendered ineffective due to other indicators on the user's screen which are competing for attention and the user's trust.

2.4.2 Studying Common Warning Failures

Using both the C-HIP model and the questions posed by Cranor [140, 31], many warning pitfalls become apparent. When a warning fails, we can use these questions to come up with explanations for the failure. Knowing how a warning is failing sheds light on creating a more effective warning. Figure 2.3 depicts each of these questions and the common ways in which a warning may fail. For instance, if a user does not notice an indicator (or the absence of an indicator), it is likely because the user either did not know to look for it or because the warning was not displayed prominently.

The C-HIP model has been applied to many differing areas in order to create more effective indicators, though all of these applications have been for static warnings in the physical world, rather than dynamic warnings used by computer software. In 2008, Cranor used the C-HIP model as a basis for the "human-in-the-loop security framework," which builds on the C-HIP model by adapting it to better fit computer security scenarios. Because the C-HIP model was created with static warnings in mind, it does not account for failures of technology, nor does it incorporate the capabilities of the human who is actually processing the warning (e.g. a warning may advise a course of action that is only comprehensible to technically savvy individuals) [32].

In his thesis, Chris Masone used the human-in-the-loop framework to analyze an email security system [88]. However, to date, I am unaware of any prior work that has applied the C-HIP model to computer security warnings using empirical data. I use the C-HIP model extensively in this thesis to qualitatively examine the ways in which trust indicators fail, by specifically examining indicators for website privacy

policies, alerting users to phishing websites, and SSL errors. The studies that I conducted helped me to create guidelines to counteract the common failures in Figure 2.3 for both critical and non-critical trust indicators.

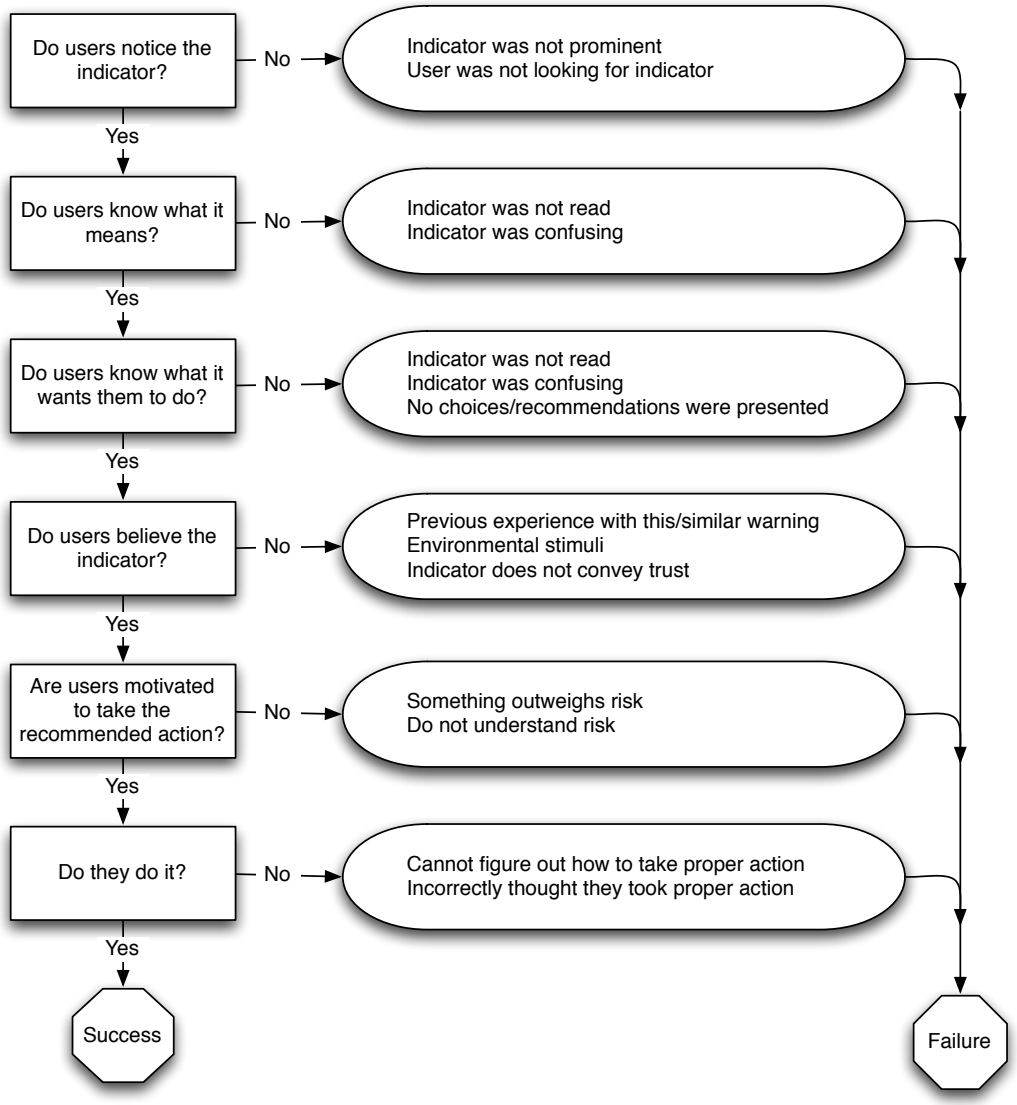


Figure 2.3: Diagram of common warning failures using the C-HIP model.

Chapter 3

Phishing Warning Study

This chapter is largely a reproduction of a paper co-authored with Lorrie Cranor and Jason Hong [48]. Thanks to the members of the Supporting Trust Decisions project for their feedback, and Matthew Williams for his assistance. This work was supported in part by the National Science Foundation under grant CCF-0524189.

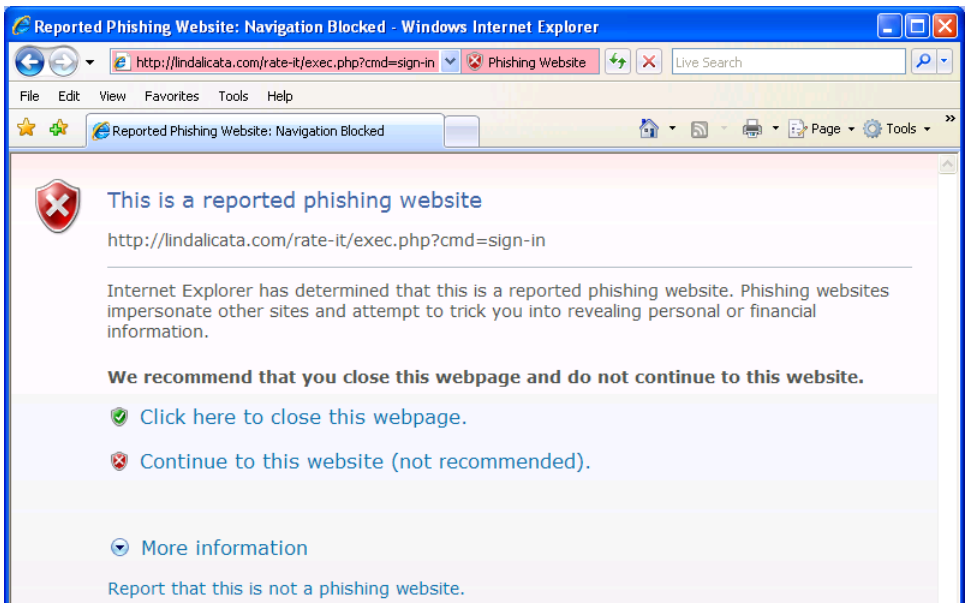


Figure 3.1: The active Internet Explorer 7.0 phishing warning.

In this study we compared the effectiveness of active and passive phishing warnings by analyzing them using a warning analysis methodology used by researchers in the warning sciences field, called the "Communication-Human Information Processing Model" (C-HIP) model [140]. The purpose of this study was to examine whether active warnings are more likely than passive warnings to help users make better trust decisions when they encounter potential phishing websites. We used the C-HIP model to examine how

Figure 3.2: The passive Internet Explorer 7.0 phishing warning.

both types of warnings can fail users in practice, and then we created recommendations for preventing these failures.

Because phishing is a serious threat that can result in substantial financial loss, newer web browsers now include phishing warnings. These warnings are "active" warnings that force users to take notice by interrupting them. Microsoft's Internet Explorer 7 includes both active and passive phishing warnings (Figures 3.1 and 3.2, respectively). When IE7 encounters a confirmed phishing website, the browser will display an active warning message giving the user the option of closing the window (recommended) or displaying the website (not recommended). This warning is a full screen error, which turns the URL bar red if the user chooses to display the website (Figure 3.1). The passive indicator, a popup dialog box, is displayed to the user when the browser believes a website is suspicious (Figure 3.2), but that website has not been verified as being a phishing website (i.e. it does not appear on a blacklist). We consider this warning to be more passive because it does not give the user any choices and it can be easily dismissed.

Firefox 2.0 also includes an active phishing warning, which was part of the Google Toolbar extension for previous versions of Firefox. When a user encounters a confirmed phishing website, a non-interactive dimmed version of the website is displayed with an overlayed dialog box. The user is given a choice between continuing to the site or leaving. The user may also click the red 'X' in the corner of the warning, which has the same effect as continuing to the website (Figure 3.3).

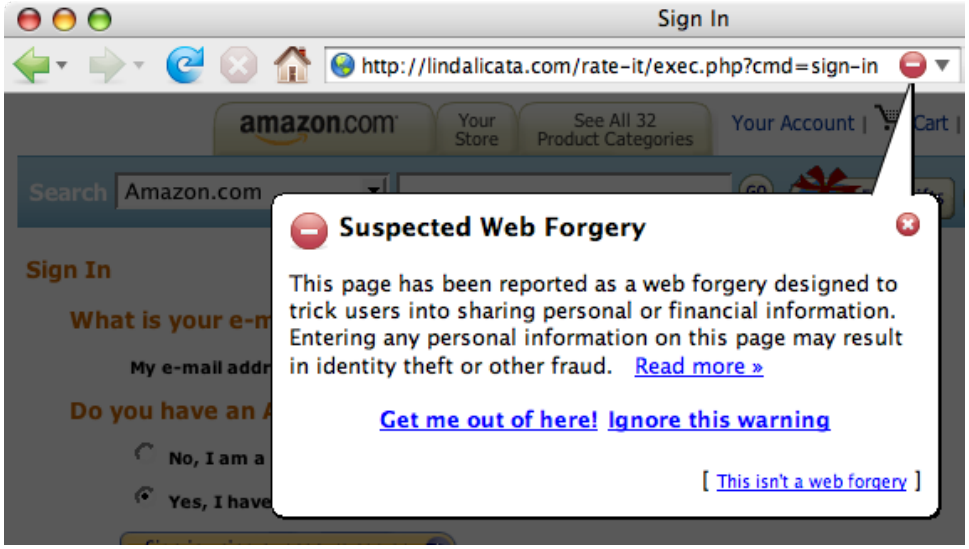


Figure 3.3: The active Firefox 2.0 phishing warning.

3.1 Methodology

The primary purpose of this study was to examine the effectiveness of phishing warnings found in current web browsers. These warnings serve as the last line of defense against a user divulging his or her sensitive information to a con artist. In other words, prior to these warnings being displayed, it is likely that users believe they are visiting legitimate websites. Thus, we needed participants to fall for the phishing messages we sent them during our study so that they would be in a similar state of mind when they encountered the warnings. At the same time, we needed our attack to be plausible. Thus, we simulated a spear phishing attack. Spear phishing "involves personalized emails or emails sent to a specifically targeted group, such as employees of a particular organization" [44]. For instance, a phisher might send a message to email addresses at *aol.com* announcing account changes impacting AOL users. Since all the recipients are AOL users, this scam may have increased credibility because the targets believe it to be relevant to them. In our study, if participants did not believe our phishing messages to be credible, they would be less likely to follow the links and thus would not see the browser warnings.

We were concerned that if participants knew the true nature of our study, their behaviors would be biased by either the Hawthorne or Milgram effects. To minimize these effects, we framed our study as an "online shopping study"—items were purchased online, and then we sent the participants phishing messages claiming to be from those shopping websites. Participants were told that we were examining how they interact with shopping websites and that they needed to think aloud during their purchases. After the first purchase was made, participants checked their email to confirm that the order was going to be shipped, thereby encountering the first phishing message. Once the participants were confident that the first purchase had been completed, instructions were provided for the second purchase. This purchase was then made using a different website, and a different phishing message was sent. Participants in the experimental conditions were given an exit survey before leaving. In this section we will provide the details of our recruitment process and the study design.

3.1.1 Recruitment

This study was designed as a between-subjects study, with four different conditions using the Internet Explorer 7.0 and Firefox 2.0 web browsers: participants were shown either the Firefox warning (Figure 3.3), the active IE warning (Figure 3.1), the passive IE warning (Figure 3.2), or no warning at all. When we performed this study in June of 2007, users of Internet Explorer and Firefox comprised 59% and 34% of all Internet users, respectively [105]. Additionally, both browsers have automatic update features. Thus, it is only a matter of time before most users will be using the newest versions of these browsers which contain active phishing warnings. We began recruiting participants in May of 2007.

We did not tell participants that we were studying online security because we wanted to simulate a natural environment by not priming them to security concerns. We recruited participants from all over Pittsburgh in order to make our results generalizable. We attached flyers to telephone posts, bus stops, and community bulletin boards. We also posted online to Craigslist and a CMU website for recruiting study participants. We constructed a screening survey to screen out technically savvy individuals, users of certain web browsers, participants in previous phishing studies, and users of certain email providers (Appendix A). We also used this survey to glean some basic demographic information from participants, such as age, gender, occupation, prior online shopping experience, etc.

Participants who contacted us after seeing a recruitment flyer were directed to our online screening survey. Since we were examining the newest versions of Firefox (2.0) and IE (7.0) to include the active warnings, we made sure that all participants in the experimental conditions already used one of these browser versions. Thus the screening survey included a question about current browser version (with graphics depicting how to determine the version) to screen out users of other web browsers.

Since our lab has conducted previous studies on phishing, we were concerned about the potential for priming of prior participants. Thus we disqualified anyone who had previously participated in a phishing-related study. We were also concerned that savvy users would not believe the emails, and thus not be exposed to the warnings. We asked four questions to gauge each participant's technical knowledge:

- Have you ever designed a website?
- Have you ever registered a domain name?
- Have you ever used SSH?
- Have you ever configured a firewall?

In our pilot we discovered that participants who answered yes to all four questions were just as likely to believe the phishing emails as all other participants. Thus, we decided not to disqualify participants based on these questions, and instead decided to use them in our analysis.

We tried to make our scenarios as realistic as possible by requiring participants to use their own email accounts and financial information for the purchases. The screening survey explicitly asked whether or not they could check their email using a web browser on a foreign computer. We also asked them to enter their email addresses so that we could contact them as well as to determine which email provider they were using. We initially found that some of the larger free email providers were detecting our phishing messages and filtering them out. We minimized this problem by implementing DKIM and SPF on our outgoing mail server to help recipient mail servers verify the message sender.[1,2]

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

[2] http://www.openspf.org/

Of the 282 individuals who completed our screening survey, only 70 qualified and showed up. Despite our efforts to screen out individuals who used email providers that were likely to filter out our messages, we still found that we could not collect data from ten participants because they did not receive either of our phishing messages. These ten participants were not included in our results.

Based on the browser versions that they indicated in the screening survey, participants were placed in one of the four conditions. The control condition, in which participants saw no warnings, was comprised of users of both browsers. The average age of participants was 28 ($\sigma = 10.58$), and there was no significant difference between the groups in terms of age or gender. The Firefox condition consisted of 20 users of Firefox 2.0, while the other two experimental conditions consisted of users of Internet Explorer 7 (20 participants in the active IE condition and 10 participants in the passive IE condition). The ten participants in the control group all used an older version of one of the two browsers. The control group was used to determine whether or not participants were willing to enter information into our phishing websites in the absence of any warning messages. This told us whether the warning was affecting phishing susceptibility or if it could be attributed to some other factor. The group sizes were chosen based on a power analysis performed prior to recruitment.

We were initially concerned that the self-selected nature of the groups (based on web browser preference) may have biased our study. However, we found no statistical differences between the average number of hours participants in each group claimed to use the Internet, nor with regard to the average number of email messages participants claimed to receive. In each of the active warning groups, exactly seven participants answered "no" to all of the questions used to gauge technical prowess. Looking at the participants who answered "yes" to all four questions, there were four in the Firefox condition, one in the active IE condition, and two in the passive IE condition. These differences were not significantly different. Thus, we have reason to believe that there were equal numbers of novices in each group.

3.1.2 Scenarios

We decided to spoof Amazon and eBay since they were the most commonly phished non-bank websites [98]. Thus, regardless of familiarity with the real websites, it is likely that participants have previously encountered phishing messages claiming to be from these websites. Our spoofed websites consisted of login forms for usernames and passwords. To make these websites look authentic, we registered two domain names: *ebay-login.net* and *amazonaccounts.net*. The websites were designed to mimic the login pages of the original websites. We created two spoof URLs at each domain name in order to trigger the two different warnings in IE, and a third that did not trigger any warnings.

We took steps to ensure our phishing websites triggered the warnings in each web browser. Firefox downloads its locally stored blacklist from Google, so we modified it locally to include our URLs [97]. Microsoft agreed to add our spoof URLs to their remote blacklists, causing those URLs to trigger the IE phishing warnings.

We copied two common phishing emails spoofing Amazon and eBay and changed the content to fit our study. The message claiming to be from Amazon was sent out in plain text and informed the recipient that the order was delayed and would be cancelled unless the recipient clicked the included URL (Figure 3.4). The message claiming to be from eBay was written in HTML and informed the recipient that all international orders needed to be confirmed by visiting a URL contained within the message (Figure 3.4). Both messages contained random order numbers to help convince the recipients of their legitimacy, though

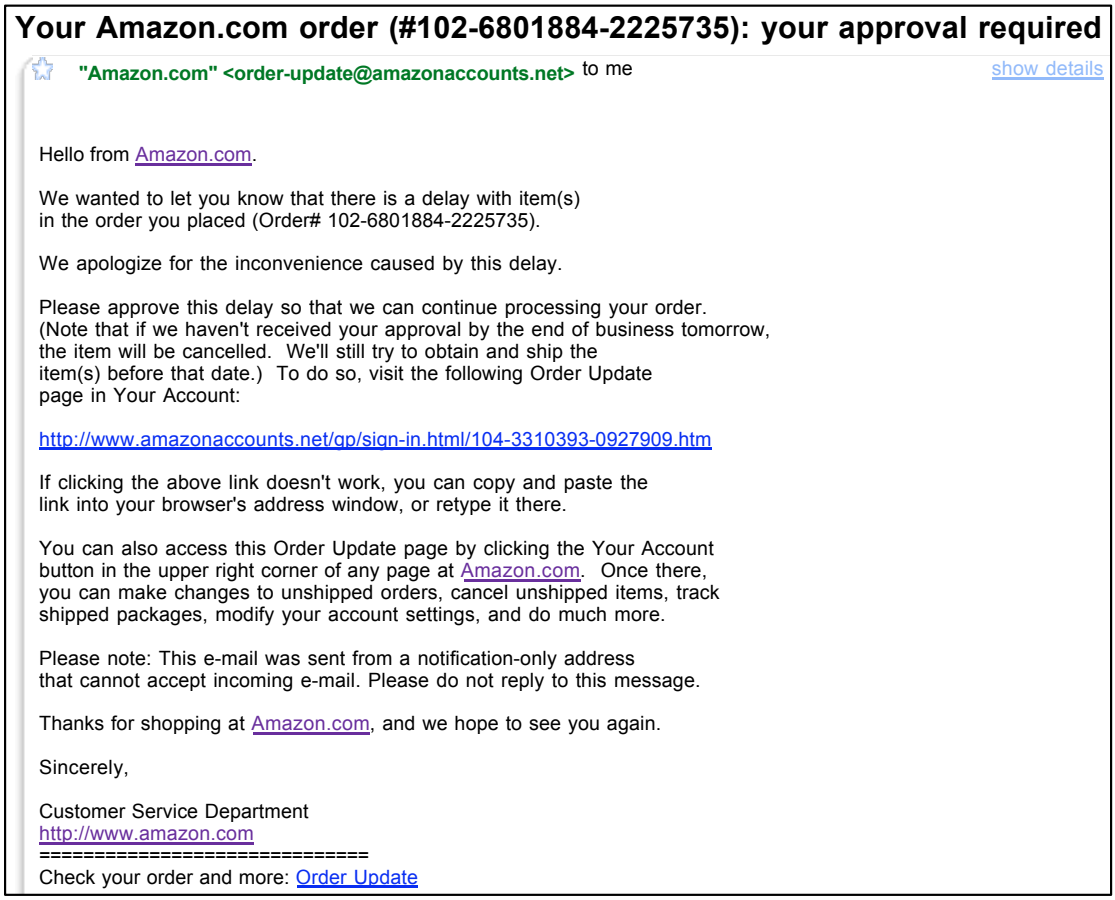


Figure 3.4: A screenshot of the phishing email that we sent claiming to be from Amazon.

no information specific to the recipients or their purchases was included in these messages in order to make our attacks realistic. The scenario was such that it would have been entirely possible for a person to have just completed a purchase from one of these websites and then received a generic phishing message spoofing that same website. It is also possible for a phisher to monitor wireless Internet traffic and conduct a similar spear phishing attack after detecting a purchase. We believe that the coincidental nature of this attack was the reason why many more participants fell for our attacks than what has been found in similar studies [44, 42, 110, 81, 112]. Previous phishing studies have spoofed companies with whom victims had relationships. However we are unaware of any user studies that have used phishing messages timed to coincide with a transaction with the spoofed brand.

Participants arrived at our laboratory and were told that they would be purchasing two items online from Amazon and eBay. We randomized the order in which the purchases were made. We also informed participants that we were recording them, so they needed to think aloud about everything they were doing. Participants did the study individually with the experimenter sitting behind them in the laboratory.

We were concerned that if we allowed participants to purchase whatever they wanted, they might take

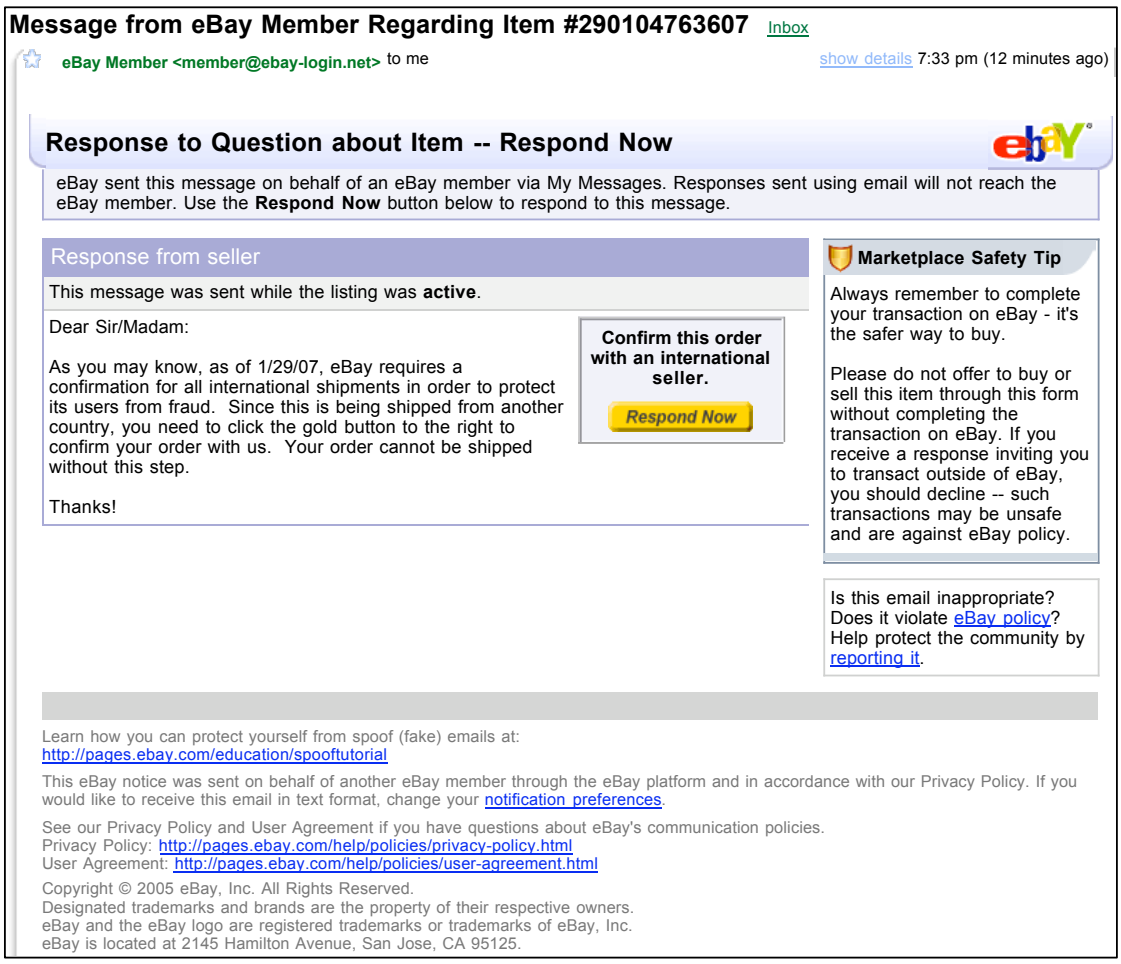


Figure 3.5: A screenshot of the phishing email that we sent claiming to be from eBay.

too long to decide, and that other factors might confound our results. We also wanted participants to focus on buying cheap items so that we could reimburse them for both purchases while still giving them enough additional money for their time. We limited the scope of the purchases by asking them to purchase a box of paper clips from Amazon, which cost roughly $0.50, plus around $6 in shipping (the exact prices changed with each order since all participants did not purchase the same type and quantity of paperclips). We asked participants to make their eBay purchases from a cheap electronics store based in Hong Kong that sold a variety of items for around $5-$10, including shipping. Participants were compensated $35 for their time and the purchases, which were made using their personal credit cards.

After each purchase, participants received a sheet of five questions relating to shopping. These questions were part of an unrelated study on shopping behaviors, but helped our study by convincing participants that they were indeed participating in a shopping study. While the participants were busy answering these questions, the experimenter sent them a phishing message claiming to be from Amazon or eBay, depending

on where the purchase was made. We constructed a web interface for the study, so that the experimenter only needed to enter an email address, the brand to spoof, and the experimental condition, which changed the URL to trigger a specific browser warning.

After the written questions were completed, the experimenter told the participant to "check your email to make sure that the order is confirmed and ready to ship so we can move on." When participants checked their email, they encountered legitimate messages relating to their orders as well as the phishing message. After examining and interacting with all of the messages, participants received a set of instructions for the second purchase. After participants checked their email after the second purchase, thereby encountering the second phishing message, an exit survey was administered (Appendix B). This online exit survey contained questions about participants' reactions to the warning messages. The experimenter observed participants fill this out and asked followup questions if any of the responses were too terse or did not seem to follow the behaviors exhibited during the experiment. Those in the control group were not asked to complete an exit survey as they had not seen any warnings. Participants took an average of forty minutes to complete all the tasks and were given $35 in cash before leaving.

We were initially concerned that since participants did not explicitly want the items, the results might be skewed in favor of participants acting more cautious. However, we believe their desire to complete the study negated this. Thus, the desire to buy the items to complete the study was likely just as strong as if the participant were at home purchasing a desired item. Additionally, we do not believe that the cost of the items played any role since an attacker could use the stolen account credentials to make any number of larger purchases. Though it is unclear if participants conceptually understood this, as we will see in the next section.

3.2 Results and Analysis

Overall we found that participants were highly susceptible to our spear phishing attack. However, users of the active phishing warnings were largely protected, since 79% chose to heed them by not entering information on the websites. We found a significant difference between the active IE and Firefox warnings ($p < 0.0004$ for Fisher's exact test) as well as no significant difference between the passive IE warning and the control group (i.e. significantly more users were helped by the active Firefox warning than the active IE warning, while the passive IE warning was not observed to be any different than not displaying any warning). We also found significant differences between the active IE warning and the control group ($p < 0.01$) demonstrating that the active IE warning is still significantly better than not displaying any warning. Table 3.1 depicts these results. In this section we examine how participants reacted to the initial phishing messages, and then we use the C-HIP model to analyze why certain warnings performed better than others.

3.2.1 Phishing Susceptibility

Our simulated spear phishing attack was highly effective: of the 106 phishing messages that reached participants' inboxes, participants clicked the URLs of 94 of them (89%). While all participants made purchases from both Amazon and eBay, not every participant received both of our phishing messages due to email filtering. Only two participants (3%) did not attempt to visit any of the phishing URLs. Of the 46 participants who received both phishing messages, 43 clicked the Amazon link and 37 clicked the eBay link.

Condition Name	Size	Clicked	Phished
Firefox	20	20 (100%)	0 (0%)
Active IE	20	19 (95%)	9 (45%)
Passive IE	10	10 (100%)	9 (90%)
Control	10	9 (90%)	9 (90%)

Table 3.1: An overview depicting the number of participants in each condition, the number who clicked at least one phishing URL, and the number who entered personal information on at least one phishing website. For instance, nine of the control group participants clicked at least one phishing URL. Of these, all nine participants entered personal information on at least one of the phishing websites.

However this difference was not statistically significant, nor were there any significant correlations based on which phishing message was viewed first. This indicates that branding likely played less of a role and the coincidental nature of the attack was therefore more responsible for tricking participants. It should also be noted that every participant in the control group who followed a link from an email message also submitted information to the phishing websites (Table 3.1). Thus, in the absence of security indicators, it is likely that this type of phishing attack could have a success rate of around 89%.

We analyzed responses to the technical ability questions in our recruiting survey mentioned in the Section 3.1.1 and noticed a negative trend between technical experience and obeying the warnings among Internet Explorer users; users with more technical experience were more likely to ignore the warnings. With Firefox, technical experience played no role: all users obeyed the warnings regardless of their technical experience.

We did not actually collect any information entered into the phishing websites. Instead the experimenter observed each participant and noted when they submitted information. Thus we cannot conclusively say whether all participants entered their correct information. However, the experimenter did note that all usernames were entered correctly, and no participants denied entering their correct information when asked in the exit survey.

We found that participants had very inaccurate mental models of phishing. Both of our phishing messages contained language that said the orders would be cancelled if they did not visit the URLs. Thirty-two percent of the participants who heeded the warnings and left the phishing websites believed that their orders would be cancelled as a result—they believed that the emails were really sent from eBay and Amazon, yet at the same time understood that they were visiting fraudulent websites not affiliated with these companies. We asked 25 of the participants how they believed the fraudulent URLs came to them, and only three recognized that the emails had been sent by someone not affiliated with either eBay or Amazon (we added this question after we had already received data from the majority of study participants). Thus, there seems to be some cognitive dissonance between recognizing a fraudulent website and the fraudulent email that linked to it. This raises grave concerns about Internet users' susceptibility to phishing. Highly targeted phishing attacks will continue to be very effective as long as users do not understand how easy it is to forge email. At the same time, effective browser warnings may mitigate the need for user education, as we will now show.

Condition Name	Sample Size	Saw Warning	Read Warning	Recognized Warning	Understood Meaning	Understood Choices
Firefox	20	20	13	4	17	19
Active IE	20	19	10	10	10	12
Passive IE	10	8	3	5	3	5

Table 3.2: This table depicts the number of participants in each experimental condition, the number who saw at least one warning, the number who completely read at least one warning, the number who recognized the warnings, the number who correctly understood the warnings, and the number who understood the choices that the warnings presented.

3.2.2 Attention Switch and Maintenance

The first stage in the C-HIP model is "attention switch." If a warning is unable to capture the user's attention, the warning will not be noticed and thus be rendered useless. Unlike the passive indicators examined by Wu et al. [146], the active warnings in Firefox and Internet Explorer get the user's attention by interrupting their task—the user is forced to choose one of the options presented by the warning.

This was not the case with the passive warning in IE (Figure 3.2). This warning is a single dialog box with only the option to dismiss it. We observed that it could take up to five seconds for this warning to appear. If a user starts typing during this period, the user's keystrokes will inadvertently dismiss the warning. Six of the ten participants in this condition never noticed the warning because their focus was on either the keyboard or the input box. Two of these participants had this happen on both phishing websites, so they had no idea they were ever exposed to any warnings. We found no statistical significance between this condition and the control group. Thus, this type of warning is effectively useless because it is so easy for it to go unnoticed.

Effective warnings must also cause attention maintenance—they must grab the users' attention long enough for them to attempt comprehension. We examined the number of participants who read the warnings (as determined by self-reporting and confirmed by the observations of the experimenter) in order to determine their effectiveness at attention maintenance. Table 3.2 shows the number of warnings read and the number of participants who claimed to have seen the warnings prior to this study, for each experimental condition.

Not counting the two participants who failed to notice the warnings entirely, and the participant in the active IE condition who did not click on the URLs, we found that twenty-six of the remaining forty-seven (55%) claimed to have completely read at least one of the warnings that were displayed. When asked, twenty-two of these twenty-six (85%) said they decided to read the warning because it appeared to warn about some sort of negative consequences.

Upon seeing the warnings, two participants in the active IE condition immediately closed the window. They went back to the emails and clicked the links, were presented with the same warnings, and then closed the windows again. They repeated this process four or five times before giving up, though never bothered to read the warnings. Both said that the websites were not working. Despite not reading or understanding the warnings, both were protected because the warnings "failed safely." Thus, if users do not read or understand

the warnings, the warnings can still be designed such that the user is likely to take the recommended action.

Nineteen participants claimed to have previously seen these particular warnings. A significantly higher proportion of participants in the active IE condition (50%) claimed to have recognized the warnings as compared to participants in the Firefox condition (20%; $p < 0.048$ for Fisher's exact test). Many of the participants who encountered the active IE warning said that they had previously seen the same warning on websites which they trusted, and thus they ignored it. It is likely that they did not read this phishing warning because IE uses a similar warning when it encounters an expired or self-signed SSL certificate. Therefore they did not notice that this was a slightly different and more serious warning.

We found a significant negative Phi correlation between participants recognizing a warning message and their willingness to read it completely ($\phi = -0.31, p < 0.03$). This implies that if a warning is recognized, a user is significantly less likely to bother to read it completely (i.e. habituation). Thus, very serious warnings should be designed differently than less serious warnings in order to increase the likelihood that users will read them. This was also the basis for Brustoloni and Villamarín-Salomón's work on dynamic warnings [19].

3.2.3 Warning Comprehension

A well-designed warning must convey a sense of danger and present suggested actions. In this study we asked participants what they believed each warning meant. Twenty-seven of the 47 participants (57%) who saw at least one of the warnings correctly said they believed that they had something to do with giving information to fraudulent websites (Table 3.2). Of the 20 participants who did not understand the meaning of the warnings, one said that she did not see it long enough to have any idea, while the others had widely varying answers. Examples include: "someone got my password," "[it] was not very serious like most window[s] warning[s]," and "there was a lot of security because the items were cheap and because they were international."

Using Fisher's exact test, we found that those using Firefox understood the meaning of the warnings significantly more than those exposed to the active IE warnings ($p < 0.041$) and the passive IE warnings ($p < 0.005$), though we found no significant difference between the active and passive IE warnings. We found a significant Phi correlation between completely reading a warning and understanding its meaning for the active IE warning ($\phi = 0.48, p < 0.037$), but not for Firefox. Since all but one Firefox user correctly understood what the warning wanted them to do, this implies that users did not need to completely read it to know the appropriate actions to take.

Overall, 31 of the 47 participants who noticed the warnings mentioned that they thought they were supposed to leave the website or refrain from entering personal information. Those who did not understand the warnings provided responses such as "panic and cancel my accounts," "confirm information about the orders," and "put in my account information so that they could track it and use it for themselves."

3.2.4 Attitudes and Beliefs

We asked participants how their attitudes and beliefs influenced their perceptions and found a highly significant correlation between trusting and obeying the warnings (i.e. users who did not trust the warnings were likely to ignore them; $\phi = 0.779, p < 0.0005$). More telling, all but three participants who ignored a warning said it was because they did not trust the warning. Two of the participants who ignored the warnings in the active IE group said they did so because they trusted them but thought the warnings were not very severe ("since it gave me the option of still proceeding to the website, I figured it couldn't be that bad"). The

other participant who trusted the warning yet ignored it was in the passive IE group and blamed habituation ("my own PC constantly bombards me with similar messages"). All three of these participants questioned the likelihood of the risks, and thus were more interested in completing the primary task.

We found a significant correlation between recognizing and ignoring a warning ($\phi = 0.436$, $p < 0.002$). This further implies that habituation was to blame when participants ignored warnings: they confused them with similar looking, but less serious warnings, and thus did not understand the level of risk that these warnings were trying to convey. This was only a problem for the warnings used by IE, as all the Firefox users obeyed the warnings (though only 20% claimed to have seen them before, compared to the 50% with IE). The IE users who ignored the warnings made comments such as:

- "Oh, I always ignore those"
- "Looked like warnings I see at work which I know to ignore"
- "Have seen this warning before and [it] was in all cases [a] false positive"
- "I've already seen such warnings pop up for some other CMU web pages as well"
- "I see them daily"
- "I thought that the warnings were some usual ones displayed by IE"

A warning should not require domain knowledge for a user to understand it. In order to examine whether prior knowledge of phishing impacted user attitudes towards the warnings, we asked them to define the term "phishing." Twenty-six of the forty-seven participants who noticed the warnings were able to correctly say they had something to do with using fraudulent websites to steal personal information. We performed a Phi correlation and found a significant correlation between knowing what phishing is and both reading ($\phi = 0.47$, $p < 0.001$) and heeding ($\phi = 0.39$, $p < 0.007$) the warnings. Thus, if a user does not understand what phishing is, they are less likely to be concerned with the consequences, and thus less likely to pay attention to the warning.

3.2.5 Motivation and Warning Behaviors

Table 3.1 depicts the number of participants from each condition who fell for at least one phishing message. Some participants only clicked on one of the two phishing messages, and in other cases some participants only received one phishing message due to email filtering.

Overall we found that active phishing warnings were significantly more effective than passive warnings ($p < 0.0002$ for Fisher's exact test). We showed the passive Internet Explorer warning to ten different participants, but only one participant heeded it and closed the website, whereas the other times participants dismissed it and submitted personal information to the phishing websites (in two of these cases participants failed to notice the warnings altogether). We found that this passive warning did not perform significantly different than the control group ($p < 1.0$ for Fisher's exact test). The active IE warning was ignored by nine participants, while in the Firefox condition every participant heeded the warning and navigated away from the phishing websites. This was a highly significant difference ($p < 0.0004$, for Fisher's exact test), however the active IE warning still performed significantly better than the control condition ($p < 0.01$) and the passive IE warning ($p < 0.044$).

Qualitatively, we examined why participants were motivated to heed or ignore the warnings. A total of thirty-one participants chose to heed the warnings, and in twenty-three of these cases participants said that the warnings made them think about risks:

- "I didn't want to get burned"
- "...it is not necessary to run the risk of letting other potentially dangerous sites to get my information"
- "I chose to heed the warning since I don't like to gamble with the little money I have"
- "I felt it better to be safe than sorry"
- "I heeded the warning because it seemed less risky than ignoring it"

Participants who chose to submit information said that they did so because they were unaware of the risks (i.e. they did not read the warnings), were used to ignoring similarly designed warnings (i.e. habituation), or they did not understand the choices that the warnings presented.

3.2.6 Environmental Stimuli

In the passive IE condition, three of the participants who ignored the warnings said they did so because they incorrectly placed some degree of trust in the phishing website because of stimuli other than the warning messages. When asked why they chose to ignore the warnings, one participant said she had "confidence in the website." Another participant ignored the warning "because I trust the website that I am doing the online purchase at." These answers corroborate Fogg's work, showing that the look and feel of a website is often the biggest trust factor [55]. Participants who ignored the active IE warning provided similar answers, and also said that they ignored the warnings because they trusted the brands that the emails had spoofed.

We also found that when some participants saw the warnings, they examined other security context information before making a decision. One Firefox user reexamined the original phishing email and noticed the lack of any personalized information. She then decided to "back out and log in from the root domain to check." After seeing the warnings, ten other Firefox users also examined either the URL bar or the email headers. Some observations included: "The URL did not match the usual eBay URL and so it could be fraudulent;" "I did look at the URL that I opened from the email, and the sender of the email, to confirm that they did look suspicious;" and "it made me look at the web address which was wrong." One participant in the passive IE condition and three in the active IE condition incorrectly used this information to fall for the phishing attacks. Some of the comments included: "The address in the browser was of amazonaccounts.com which is a genuine address" and "I looked at the URL and it looked okay."

Finally, at least four participants claimed that the timing of the phishing emails with the purchases contributed to them ignoring the warnings. It is unclear how susceptible these participants would have been to a broader phishing attack, rather than the targeted attack that we examined.

3.3 Discussion

In this section we provide some recommendations for improving the design of phishing indicators based on the results of our study.

Interrupting the primary task — Phishing indicators need to be designed to interrupt the user's task. We found that the passive indicator, which did not interrupt the user's task and therefore easily went unnoticed by participants, was not significantly different than not providing any warning . The active warnings were effective because they facilitated attention switch and maintenance by interrupting the participants' tasks, therefore forcing them to attempt to comprehend the warning enough to determine how to complete the task.

Preventing habituation — Phishing indicators need to be distinguishable from less serious warnings and used only when there is a clear danger. Even if a warning is able to capture and maintain attention, it

may fail if users confuse it with a less serious warning and therefore fail to comprehend it. In this study, users ignored the passive warnings because they looked like many other warnings that users have ignored without consequences, thus they appear to be "crying wolf." Even the active Internet Explorer warning was not read in a few cases because users mistook it for other less-serious IE warnings. More people read the Firefox warnings because they are designed unlike any other warnings. Dynamic warning messages may help prevent habituation [19], but most importantly, high-risk warnings should be instantly distinguishable from low-risk warnings.

Clearly state risk and consequences — If users notice a warning and understand what it says, it still may fail if they do not believe they are in any danger. In this study we saw that users who knew what phishing was were significantly more likely to obey the warnings, whereas users who did not understand the risks were likely to ignore the warnings. Errors in the *Attitudes and Beliefs* stage of the C-HIP model may be minimized by clearly stating the risks and consequences of ignoring the warning in language that all users are likely to understand.

Providing clear choices — Phishing indicators need to provide the user with clear options on how to proceed, rather than simply displaying a block of text. Even if the text recommends a course of action, users are unlikely to follow it unless they are provided with a means, such as a button or link. The users that noticed the passive Internet Explorer warning, read it but ignored it because they did not understand what they were supposed to do. They understood it had something to do with security, but they did not know how to proceed, so they therefore proceeded with their task and fell for the attack. In contrast, the active warnings presented choices and recommendations which were heeded by most participants. Wu found similar results with regard to providing users with clear choices [145].

Failing safely — Phishing indicators must be designed such that one can only proceed to the phishing website after reading the warning message. Users of the active Internet Explorer warning who did not read the warning or choices could only close the window to get rid of the message. This prevented them from accessing the page without reviewing the warning's recommendations. However, users of the passive Internet Explorer warning had the option of clicking the familiar 'X' in the corner to dismiss it without reading it, and accessing the page anyway.

Altering the phishing website — Phishing indicators need to distort the look and feel of the website such that the user does not place trust in it. This can be accomplished by altering its look or simply not displaying it at all. The overall look and feel of a website is usually the primary factor when users make trust decisions [55]. When the website was displayed alongside the passive indicators, users ignored the warnings because they said that they trusted the look of the website.

3.4 Conclusion

This study has given us insights into creating effective security indicators within the context of phishing. Such indicators are clearly needed as 97% of participants believed the phishing emails enough to visit the URLs. Of the participants who saw the active warnings, 79% chose to heed them and close the phishing websites, whereas only 13% of those who saw the passive warnings obeyed them. Without the active warning indicators, it is likely that most participants would have entered personal information. However, the active indicators did not perform equally: the indicators used by Firefox performed significantly better than the active warnings used by IE, though both performed significantly better than the passive IE warnings (which was not significantly different from not showing any warnings in the control group).

As phishing attacks continue to evolve, it is likely that highly targeted attacks will become more prevalent. Future indicators within the phishing context need to be designed such that they interrupt the user's primary task, clearly convey the recommended actions to take, fail in a secure manner if the user does not understand or ignores them, draw trust away from the suspected phishing website, and prevent the user from becoming habituated.

Chapter 4

Warning Options Study

This chapter describes joint work with Stuart Schechter, which was previously unpublished.

In this chapter I present results of a study where we validated some of my recommendations for improved web browser phishing warnings. Specifically, we examined the role that option text has on users' trust decisions. Recall that an active warning presents users with multiple options on how to proceed, usually with one option that is recommended because it is safer (e.g. closing the web browser), and one that is not recommended (e.g. continuing to the website despite the warning). Microsoft updated their phishing warnings in the beta version of Internet Explorer 8 (IE8) to use new option text (Figure 4.1) [84]. When a user encounters a website that is suspected of being a phishing scam, the user is now advised to "go to my homepage instead." We were concerned that users may see this message and choose to proceed to the website despite the warning because the recommended option does not conceptually allow them to complete their primary tasks; they were unlikely attempting to access their homepages before seeing the warning, and therefore would be unlikely to do so after seeing the warning.

We tested how option text impacts decisions in the laboratory by creating an experimental condition that appeared to be more likely to aid in completing the primary task: "search for the real website." We also believed that this text would emphasize the threat model: they were visiting a fraudulent website designed to look like a legitimate one. In addition to examining the option text, we wanted to validate my recommendation to design the phishing warning differently from less-severe warnings in order to minimize habituation, which I had proposed in my previous study [48]. Thus, we varied both the option text and background color.

4.1 Methodology

In the Fall of 2008 we conducted a laboratory study to examine the roles of option text and arousal strength on phishing warnings. We conducted a phishing study similar in methodology to the one conducted in Chapter 3, with the addition of an eye tracker to gauge the parts of the warnings that participants were examining. Participants arrived at our laboratory, were randomly assigned to one of three experimental conditions, and then were given tasks that required them to check their email. After one of these tasks, a phishing message appeared in their inboxes, and we observed whether they clicked on it, and whether they were protected by the warning messages. After the tasks, the participants filled out an exit survey.

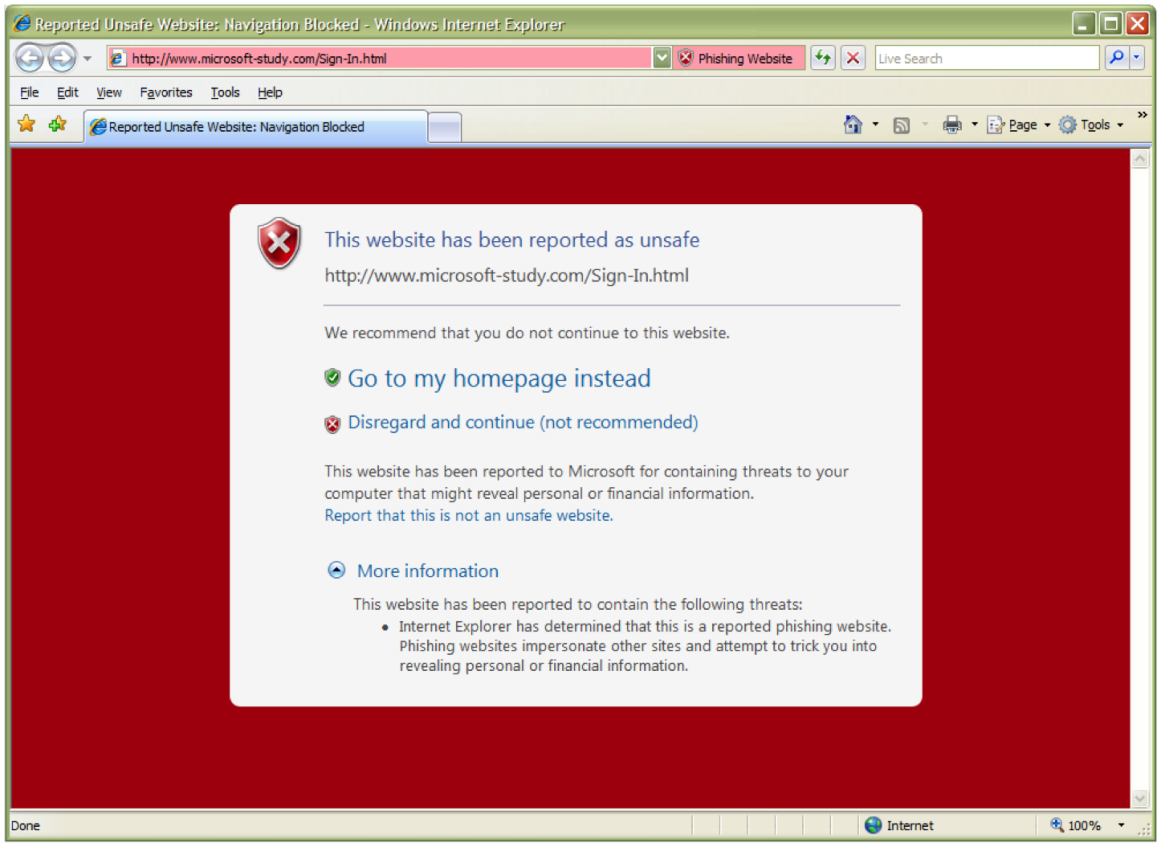


Figure 4.1: The new Internet Explorer 8 phishing warning.

4.1.1 Conditions

We created two between-group conditions to examine the role of the option text: one condition displayed a warning that recommended users "go to my homepage instead," while the warning in the other condition recommended that they "search for the real website." We will refer to these as the *home* and *search* conditions, respectively. Our hypothesis was that study participants would be less likely to heed the recommendations of the phishing warnings if those recommendations appeared unlikely to help complete a primary task; we believed that fewer participants would obey the warnings if the warnings recommended that they visit their homepages when that was not what the participants were trying to do prior to seeing the warnings. We believed that the text "search for the real website" helped to both complete a primary task as well as underscore the threat model: the website that they are being warned about is likely fraudulent and mimicking a legitimate one.

We created a third between-group condition to examine the role of habituation by removing the red border from the warning, and replacing it with a white border, so that it would look more similar to the IE7 warnings (Figure 4.2). To control for both the text and color of this warning, we set the recommended option to "go to my homepage instead." We refer to this condition as the *white* condition.

IE stores all of its full-screen security warnings as HTML within a DLL resource file, *ieframe.dll.mui* [107].

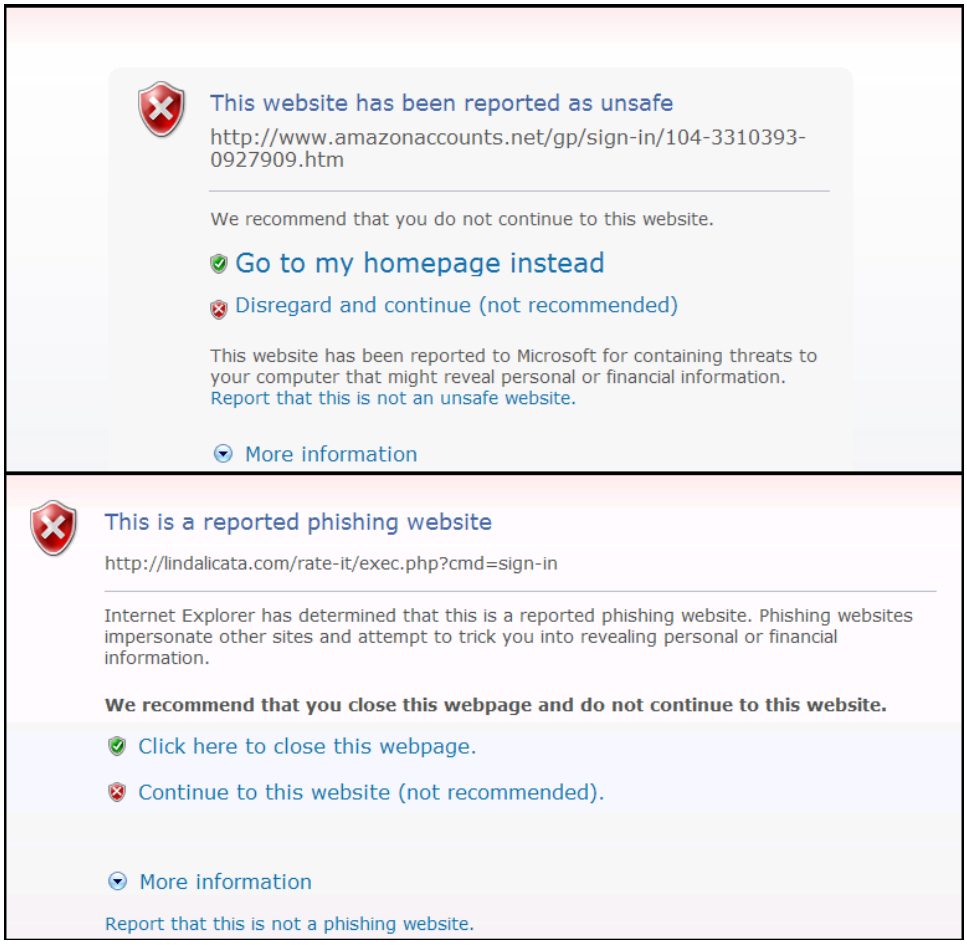


Figure 4.2: Our experimental warning condition with the white border (top), which was designed to appear similarly to the IE7 phishing warning (bottom).

To configure the computer for each condition, we created three separate versions of this resource file by modifying the HTML of the phishing warning. Prior to each experimental session, we manually switched the appropriate files to configure the computer for the appropriate condition.

4.1.2 Recruitment

We recruited 59 participants to show up at our usability laboratory during September of 2008. We selected participants who had opted in to being contacted by Microsoft to participate in user studies, and screened out participants who either did not use Hotmail for their email or did not use IE as their primary web browser. Because we were only interested in participants who were most vulnerable to phishing attacks, we screened out participants who had technical jobs, and therefore may have been less likely to fall for phishing attacks.

Participants were scheduled for individual one-hour sessions. When a participant arrived, we greeted him in the lobby of our building and then escorted him to our usability laboratory. The participant completed

a consent form and then was handed an instruction sheet. The experimenter read the instructions aloud to ensure that each participant was given the same information as all the previous participants. Before beginning the study tasks, the experimenter calibrated the eye tracker and started the screen capture program. Once the participant was ready to begin, the experimenter left the room so as to not influence the participant's behaviors.

4.1.3 Tasks

Because we were examining phishing warnings, the last line of defense before visiting a phishing website, participants needed to be in a similar state of mind as if they were seeing these warnings on their home computers. Likewise, since security is rarely a primary task (e.g. users do not sit down at the computer to "not get phished"), we needed to come up with a study design that masked the real purpose. We decided to tell participants that we were working on improvements for Windows Live Hotmail, and therefore we would be observing them interacting with their email. As an incentive, we offered them a dollar for every message that they opened during the study, and an additional four dollars if they "interacted with that message" in any way. So as to minimize the Milgram Effect—we were worried they may have felt compelled to click links in every email message in order to get paid—we told them that filing messages away or deleting them would count as an interaction.

We told participants that they would be using their actual email accounts for the study, and therefore they should behave just as they do when checking email at home. We also told them that because we cannot expect that they will receive real email during the study period, the experimenter would send them a message every ten minutes. At this point, the experimenter observed them log into their Hotmail accounts, and then left the room to observe them from our observation room.

After ten minutes had elapsed, the experimenter sent the first email message. This message was a personal message written in plaintext that asked the participant to visit *www.fandango.com* and respond with the movie they most want to see. After the next ten minutes had elapsed, the experimenter sent a second message. The second message was an HTML-based message that came from Windows Live SkyDrive[1] and invited the participant to view a shared photo album that the experimenter had posted. These two tasks were created purely for subterfuge; we wanted to convince the participants that we were only interested in how they interacted with their email.

Two minutes after participants viewed the second email message, the experimenter sent a third message that was designed to be indistinguishable from a phishing message. This message was sent from a domain other than *microsoft.com*, though claimed to be from Microsoft and encouraged readers to click a link and enter personal information on the resulting website. The domain used for the destination URL as well as sending the email, *microsoft-study.com*, was a domain that we registered for the study, which we then added to IE's phishing blacklist (thereby triggering IE's phishing warning). We sent this message outside of the ten minute interval in order to create plausible deniability; if participants knew for sure that it was sent by us as part of the study then we would not be realistically simulating a phishing attack. The message claimed to come from Windows Live, and offered participants the opportunity to enter a prize drawing if they visited the included URL. Upon arriving at this URL, participants saw one of the three warnings that we described in Section 4.1.1. If they chose to ignore the warning and proceed to the website, they were presented with a

[1] http://skydrive.live.com/

login form that appeared identical to the Windows Live login screen (except that no credentials were actually transmitted; Figure 4.3).

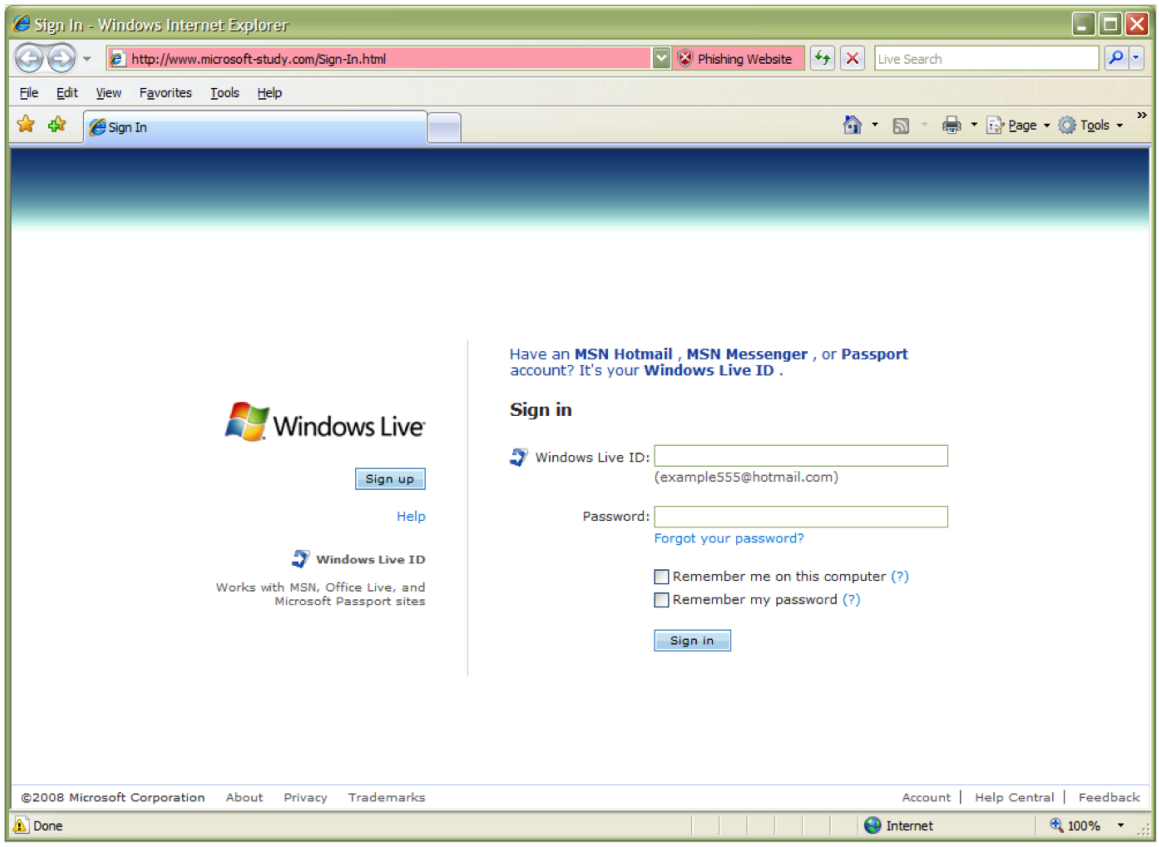


Figure 4.3: Screenshot of the destination "phishing" website that we hosted at *microsoft-study.com*. This website was designed to mimic the Windows Live login screen.

We observed participants to determine how many of them read the phishing message, how many followed the link to the website, and then whether they obeyed the warnings. After interacting with the phishing message, the experimenter returned to give the participant a written exit survey. This exit survey was designed to gather qualitative data on why participants chose to heed or ignore the warnings, as well as to gather demographic data. Upon completing the exit survey, we gave participants a voucher for their choice of Microsoft software as a gratuity.

4.2 Results

Of our 59 participants, we observed 48 who followed the links to *microsoft-study.com*.[2] This indicates that our phishing attack fooled over 80% of our study participants. Due to technical difficulties, three of these participants did not view any of the phishing warnings and therefore proceeded to enter their personal information (i.e. everyone who did not see a warning entered their information). At the same time, only 12 of the 45 (27%) participants who viewed the warnings entered personal information, whereas everyone else understood that it was safer to heed the warning.

Of the 45 participants who viewed the phishing warnings, 23 were female and 22 were male. The average age was 38.87 ($\sigma = 12.32$). We did not observe any statistically significant correlations based on age or gender. The participants who viewed phishing warnings were divided evenly among the three experimental conditions, such that there were fifteen participants per condition. Unfortunately, we found no significant differences between the conditions based on whether participants chose to heed or ignore the warnings. Additionally, we were only able to get eye tracking data from 27 of the participants, and therefore did not have enough data to use for a statistical analysis. However, we did observe a strong interaction effect based on both the red border and the text of the warnings. In the next section we present our results in terms of the effects of the red background and then the varied option text.

4.2.1 Background Color

In two of the experimental conditions, the warnings were surrounded by red borders. The purpose of this red border was to make the warnings instantly distinguishable from other less serious IE warnings. While we did not observe any differences between the conditions with regard to whether participants were more likely to get phished, we did observe that the red border caused participants to view the warnings for longer. Table 4.1 depicts the total time participants in each condition spent viewing the phishing warnings, the number of times they revisited the phishing warnings, and the average time spent viewing the warnings.[3]

We performed an ANOVA and found that participants in the *search* condition viewed the phishing warnings for a significantly longer amount of total time ($F_{2,41} = 4.754$, $p < 0.014$). Upon performing post-hoc analysis using Tukey's HSD test, we found that this was due to significant differences between the *search* and *white* conditions ($p < 0.013$), and that there were no observable differences between any of the other conditions. Likewise, when examining the total number of times that participants viewed the warnings, we found that those in the *search* condition went back to review the warning significantly more often (i.e. they closed the warning, went back to their email, reread the message, clicked the link again, etc.; $F_{2,41} = 5.046$, $p < 0.011$). This was attributed to the contrast with the *white* condition ($p < 0.012$), and to a lesser extent the *home* condition ($p < 0.061$). This shows that there appears to be an interaction effect between the background color and the option text; participants spent significantly longer analyzing the warnings only when both these features were changed. Thus, the red background likely increased arousal strength and overcame habituation effects that we observed in the previous IE phishing warnings [48]. Increasing arousal strength

[2] Of the eleven participants who did not visit the URL, three never saw the phishing messages due to technical errors (i.e. the email was never received or automatically classified as spam), while eight deleted the messages without visiting the URL because they perceived it—correctly—as being spam or part of a scam.

[3] We removed data from one participant in the *white* group after he—against directions—asked the experimenter for help and then waited for the experimenter to respond from the observation room, therefore artificially increasing the amount of time he spent viewing the warning.

Condition Name	Total Time	Average Views	Average Time
White	12.00s	1.36	9.76s
Home	17.81s	1.67	10.76s
Search	30.97s	2.67	11.84s

Table 4.1: This table shows the three experimental conditions as well as the total amount of time participants spent viewing the warnings (averaged over each condition), the average number of times participants viewed the warnings, and the average amount of time participants spent with each viewing (averaged over each condition). Participants in the *search* condition viewed the warnings significantly more frequently as well as for significantly longer periods of time in total.

prompts participants to take greater notice of warnings, which prevents errors at the *Attention Switch* and *Attention Maintenance* stages of the C-HIP model.

To validate whether the red background had distinguished the *search* and *home* conditions from the previous IE warnings, we used our exit survey to ask participants whether they had seen these particular warnings before. We found a highly significant correlation between recognizing the warnings and falling for the phishing attacks ($p < 0.002$ for Fisher's exact test); nine of the twelve "victims" said they recognized the warnings. We also found that 53% of those in the *white* condition said they recognized the warnings, as opposed to 33% and 20% in the *home* and *search* conditions, respectively. Unfortunately, these differences were not statistically significant among our sample size, though they may indicate a trend.

4.2.2 Option Text

We changed the option text in the *search* condition to "search for the real website" in order to better convey the threat posed by phishing. The purpose of this was to prevent errors in the *Comprehension/Memory* stage of the C-HIP model: if users are unlikely to understand what the warning means or is recommending that they do, then they are unlikely to obey it. We asked participants to quantify the likelihood of something bad happening when ignoring the warnings by using a 5-point Likert scale. We discovered no significant differences between the three conditions ($\mu = 3.18$, $\sigma = 1.07$), nor when we grouped the *white* and *home* conditions together so that we were just comparing the two different sets of option text.

To further examine participants' motivations for their behaviors, we asked them to rank five factors that they used in their decisions using a 5-point Likert scale: the text of the warnings, the color of the warnings, the choices the warnings presented, the URL of the destination website, and the design of the destination website. The only significant difference we observed between conditions was with regard to the destination URL (i.e. *microsoft-study.com*; $F_{2,42} = 4.469$, $p < 0.017$). Participants who were in the *search* condition were significantly less likely to say that the destination URL was a factor in their decisions. To verify that this was due to the option text, we combined the *white* and *home* conditions ($\mu = 3.33$, $\sigma = 2.02$) and compared them with the *search* condition ($\mu = 1.53$, $\sigma = 1.81$) using a t-test. We found significant differences ($t_{43} = 2.911$, $p < 0.006$), even after using the Bonferroni method to correct for multiple testing ($\alpha = 0.01$).

We asked participants to report the most important factor of the five that they ranked. Overall, there were nine participants who claimed that the URL was biggest factor in their decision of whether or not to ignore the warning. Four of these participants were in the *white* condition, three were in the *home* condition,

and two were in the *search* condition. More importantly, five of the seven participants in the *white* and *home* conditions fell for the phishing attacks. Likewise, of all the participants who fell for the phishing attack, a plurality mentioned the URL as the greatest factor in their decisions to ignore the warnings.

4.3 Discussion

Overall when it came to the resulting behaviors, we were surprised that there were fewer differences between the conditions than we expected: there were no observable differences between the number of participants who were phished in each condition. Despite marked improvements at capturing users' attentions long enough to get them to notice the warnings, and increasing the amount of time they spent noticing the warnings, a third of the participants still ultimately succumbed to the attack. We found no correlation between falling for the attack and the amount of time or frequency of times that participants viewed the warnings. Thus, while participants in the *search* condition paid more attention to the warnings, they ultimately were just as likely to take the unsafe action. We believe this occurred because we only examined the roles of the option text and background colors; participants ignored the warnings because of failures at other stages of the C-HIP model that we did not address in this study. In this section we discuss some possible reasons for why the warnings failed and how further changes may improve warning effectiveness.

4.3.1 Understanding Risks and Consequences

We asked participants what they believed the warnings wanted them to do, and we found that of the twelve who were phished, only one did not say something along the lines of "do not visit the website." This one participant, who was in the *white* condition, responded *"check the sender or link to make sure it would not be harmful."* We believe that many participants were tricked into ignoring the warnings and entered their personal information because they did not properly understand the risks and possible consequences of their actions. We asked participants to explain the danger of ignoring the warnings and coded their answers based on whether or not they understood that phishing scams attempt to steal personal information in order to commit theft. We found that only 31% understood this (14 of 45). Of the remaining 31 participants, all of them either mentioned a combination of common security threats or a single irrelevant security threat (e.g. malware):

- *"I could potentially get a virus or spyware"*
- *"Getting a virus ruining your computer"*
- *"may get malware, give out sensitive info by mistake, virus"*
- *"a possible virus that I have protection for"*
- *"Will get some spyware"*

Three of the participants who were phished (25% of 12) in our study said that they decided to ignore the warnings because they were not using their own personal computers and therefore did not care if our computer was infected with a virus. Likewise, ten of the participants who were phished (83% of 12) in our study said that they decided to ignore the warning because they knew that they were visiting a genuine

Microsoft websites (e.g. *"I disregarded it [the warning] because I saw Microsoft-study.com as the email"*). Thus, their misunderstanding of the threat model led to them disclosing their Hotmail credentials. Conveying risks better within the warnings may prevent this sorts of errors in the future.

4.4 Conclusion

In this study we showed that distinguishing severe warnings from other less-severe warnings may aid in capturing users' attentions and minimizing habituation effects. In our particular study, we used a red border to differentiate the warnings in two of our study conditions from the previous IE warnings. Overall, we found that participants spent significantly longer viewing these warnings and were less likely to say that they had seen them before. Thus, designing warnings differently based on the danger they represent may prevent errors at the *Comprehension/Memory* and *Attitudes & Beliefs* stages of the C-HIP model by decreasing the chances that users confuse them with less-severe warnings to which they may already be habituated.

We found that the role of the option text was more subtle: when we used text that emphasized the threat of visiting a fraudulent website—"search for the real website"—participants were less likely to be tricked by the URL. Whereas those participants who were given the option to "go to my homepage" were more likely to confuse our fake website, *microsoft-study.com*, with a legitimate Microsoft website. However, this by itself was not enough to prevent participants from being phished. In future warnings, designers should highlight the risks and consequences of the warnings so that users are more likely to believe that the warnings are relevant to them. If users do not believe that the warnings are relevant to them, they are more likely to ignore them and continue to malicious websites. This represents a failure at the *Motivation* stage of the C-HIP model, and can be minimized by paying attention to the rest of the text on the warning, beyond just the options.

In the next chapter of this thesis, I validate these findings by testing a new warning for SSL errors. This new warning uses descriptive text to clearly convey the threat model and potential consequences of ignoring the warning. We performed a laboratory study and discovered that when viewing the new text, participants has significantly greater risk perceptions than they did when viewing previous SSL warnings. Thus, they became motivated to perform the correct action.

Chapter 5

SSL Warning Study

This chapter is largely a reproduction of a paper co-authored with Joshua Sunshina, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor [122]. Thanks to Dhruv Mohindra, Amit Bhan, and Stuart Schechter for their assistance. This work was supported in part by Microsoft Research and by the National Science Foundation under Grants No. 0524189 and 0831428

In the two previous chapters I examined flaws in web browser phishing warnings and made recommendations on how similar web browser security warnings could be improved. In this chapter I present a new warning design used for SSL errors and present the results of a user study that I performed to validate my findings.

The warnings science literature suggests that warnings should be used only as a last resort when it is not possible to eliminate or guard against a hazard. When warnings are used, it is important that they communicate clearly about the risk and provide straightforward instructions for avoiding the hazard [119, 139]. In this paper we examine user reactions to five different SSL[1] warnings embodying three strategies: make it difficult for users to override the warning, clearly explain the potential danger facing users, and ask a question users can answer. By making it difficult for users to override the warning and proceed to a potentially dangerous website, the warning may effectively act as a guard against the hazard, similarly to the way a fence protects people from falling into a hole. While some people may still climb the fence, this requires extra effort. By clearly explaining the potential danger, warnings communicate about risk and help users to make decisions based on knowledge of the potential consequences. Finally, by asking users question they can answer, warnings are able to instruct users in the appropriate steps necessary to avoid the hazard in a given situation.

We conducted a survey of 409 Internet users' reactions to the current web browser SSL warnings. We analyzed the results to determine whether users understood the current warnings, whether they believed the risk of ignoring a particular warning varied based on the destination website, whether they believed one particular SSL error was more severe than another, and whether these responses changed based on each respondent's level of computer security expertise. We found that participants who understood the risk associated with the warnings were more likely than those who did not understand the risk to indicate that they would refrain from visiting sites with warnings that they considered risky. However, those who

[1] In this chapter we use the common convention of using the term "SSL" to refer to both the SSL and TLS protocols.

understood the risk also perceived some common SSL warnings as not very risky, and were more likely to override those warnings.

We followed up this survey with a between-subjects laboratory experiment involving 100 participants who encountered SSL warnings on an online banking website that requested their credentials and a library website that did not request any credentials. We tested the Firefox 2 (FF2), Firefox 3 (FF3), and Microsoft Internet Explorer 7 (IE7) SSL warnings. We also tested two new warnings designed to take advantage of the lessons we learned in the survey. The first warning was designed with risk in mind: it succinctly explained the risks and consequences of proceeding to the website. The second warning was context sensitive: it appeared to be more severe when the participants visited websites that required them to enter personal data. We found that most participants ignored the FF2 and IE7 warnings on both websites. Many participants who used FF3 were unable to override that warning and were thus prevented from visiting both websites. Finally, we found that participants who viewed our redesigned warnings better understood the risks and made their decisions based on the type of website they were visiting. However, despite the fact that the warnings we examined embodied the best techniques available, none of the warnings provided adequate protection against man-in-the-middle attacks. Our results suggest that, while warnings can be improved, a better approach may be to minimize the use of SSL warnings altogether by blocking users from making unsafe connections and eliminating warnings in benign situations.

5.1 SSL Survey

5.1.1 Methodology

In the summer of 2008 we conducted an online survey of Internet users from around the world to determine how they perceived the current web browser SSL warnings. We used screenshots of the warnings from FF2, FF3, and IE7. Respondents viewed screenshots of three different SSL warnings from the browser that they were using at the time they took the survey and were asked several questions about each warning (Appendix E). These questions were followed by a series of questions to determine demographic information.[2]

We showed participants warnings for expired certificates, certificates with an unknown issuer, and certificates with mismatched domain names.[3] Each warning was shown on a separate page along with its associated questions, and the order of the three pages was randomized. We included a between group condition to see if context played a role in users' responses: half the participants were shown a location bar for *craigslist.org*—an anonymous forum unlikely to collect personal information—and the other half were shown a location bar for *amazon.com*—a large online retailer likely to collect personal and financial information. We hypothesized that respondents might be more apprehensive about ignoring the warning on a website that was likely to collect personal information. Below each warning screenshot, participants were asked a series of questions:

[2] Users of web browsers other than FF2, FF3, or IE7 were only asked the demographic questions.

[3] We examined these three warnings in particular because we believed them to be the most common.

If you saw this message, would you attempt to continue to the website?
What do you believe this message means?
Have you seen this particular message before?
How likely is it that something bad would happen if you continued on to the website?
If something bad did happen from continuing on to the website, how bad do you think it would be?
What do you believe are the possible consequences of ignoring this message?

We were also interested in determining how computer security experts would respond to our survey, and if the experts' answers would differ from everyone else's answers. In order to qualify respondents as experts, we asked them a series of five questions to determine whether they had a degree in an IT-related field, computer security job experience or course work, knowledge of a programming language,[4] and whether they had attended a computer security conference in the past two years.

We recruited participants from Craigslist and several contest-related bulletin boards, offering a gift certificate drawing as an incentive to complete the survey. We received 615 responses; however we used data from only the 409 respondents who were using one of the three web browsers under study.

5.1.2 Analysis

The 409 survey responses were split up as follows: 96 (23%) used FF2, 117 (29%) used FF3, and 196 (48%) used IE7. While age and gender were not significant predictors of responses,[5] it should be noted that 66% of our respondents were female, significantly more males used FF3 ($\chi^2_2 = 34.01$, $p < 0.0005$), and that IE7 users were significantly older ($F_{2,405} = 19.694$, $p < 0.0005$). For these reasons and because respondents self-selected their web browsers, we analyzed the responses for each of the web browsers separately.

We found few differences in responses based on the type of website being visited. We found that respondents' abilities to correctly explain each warning was a predictor of behavior, though not in the way we expected: respondents who understood the domain mismatch warnings were less likely to proceed whereas we observed the opposite effect for the expired certificate warnings. This suggests that participants who understood the warnings viewed the expired certificate warnings as low risk. Finally, we found that risk perceptions were a leading factor in respondents' decisions and that many respondents—regardless of expertise—did not understand the current warnings or associated risks. In this section we provide a detailed analysis of our results in terms of warning comprehension and risk perceptions, the role of context, and the role of expertise.

Comprehension and Risk Perceptions

We were primarily interested in whether respondents would continue to the destination website if they saw a given warning. As shown in Figure 5.1, less than half the participants claimed they would continue.

We expected to see differences in behavior for each of the three types of warnings. In order for this to be the case, participants needed to be able to distinguish each of the three warnings. We asked them to explain what they thought each warning meant and coded the answers in terms of whether or not they were

[4] Respondents also specified known programming languages so that we could verify their responses.

[5] All statistics were evaluated with α=0.05. Those for which we only show a p-value were conducted using a Fisher's exact test.

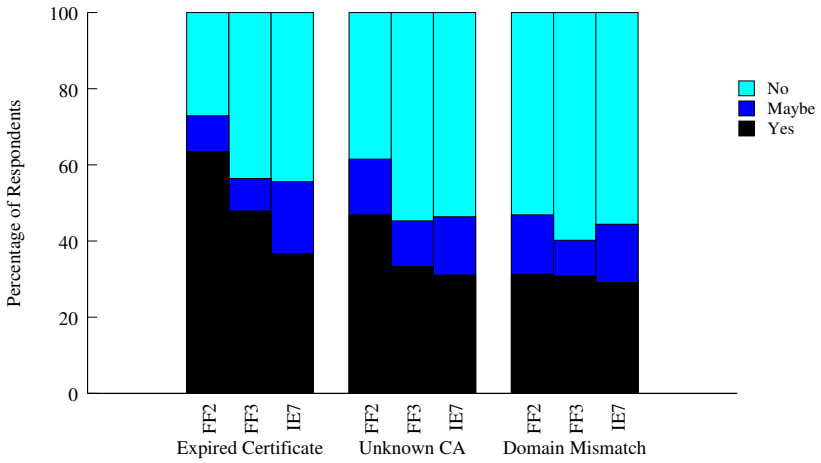


Figure 5.1: Participant responses to the question: *If you saw this message, would you attempt to continue to the website?* Because of few significant differences based on the type of website they were viewing, we combined the two conditions for this analysis.

Browser	Understood	Expired Certificate		Ignored		Unknown CA		Ignored		Domain Mismatch		Ignored	
FF2	Y	48	50%	71%		37	39%	43%		57	59%	19%	$\chi^2_2 = 9.40$
	N	48	50%	56%		59	61%	49%		39	41%	49%	$p < 0.009$
FF3	Y	55	47%	64%	$\chi^2_2 = 21.05$	35	30%	31%		46	39%	15%	$\chi^2_2 = 8.65$
	N	62	53%	34%	$p < 0.0005$	82	70%	34%		71	61%	41%	$p < 0.013$
IE7	Y	45	23%	53%	$\chi^2_2 = 11.81$	44	22%	27%		62	32%	16%	$\chi^2_2 = 7.50$
	N	151	77%	32%	$p < 0.003$	152	78%	32%		134	68%	35%	$p < 0.024$

Table 5.1: Participants from each condition who could correctly identify each warning, and of those, how many said they would continue to the website. Differences in comprehension within each browser condition were statistically significant (FF2: $Q_2 = 10.945$, $p < 0.004$; FF3: $Q_2 = 11.358$, $p < 0.003$; IE7: $Q_2 = 9.903$, $p < 0.007$). For each browser condition, the first line depicts the respondents who could correctly define the warnings, while the second depicts those who could not. There were no statistically significant differences between correctly understanding the unknown CA warning and whether they chose to ignore it.

correct.[6] As shown in Table 5.1, we discovered that FF2 users were significantly more likely to understand the domain mismatch warnings, while FF3 users were significantly more likely to understand the expired certificate warnings.

We explored warning comprehension further by examining whether those who understood the meaning of the warnings were more likely to heed or ignore them. In general, we found that users who understood the warnings tended to behave differently than those who did not. Across all three browsers, users who understood the domain mismatch warning were more likely to say they would heed that warning than users who did not understand it. In addition, FF3 and IE7 users who understood the expired certificate warnings were more likely to indicate that they would ignore these warnings and proceed to the destination website. These results are detailed in Table 5.1 and indicate that users likely perceive less risk when encountering an expired certificate, and therefore are likely to proceed. However, when encountering a domain mismatch warning, knowledgeable users perceive greater risk and are likely to discontinue.

The three warnings that we examined are displayed when the authenticity of the destination website's SSL certificate cannot be guaranteed. While each of these warnings represents a different underlying error, they represent the same threat: the user may not be communicating with the intended website or a third party may be able to eavesdrop on her traffic. In both cases, sensitive information may be at risk (e.g. billing information when performing an online purchase). In order to determine whether or not respondents understood the threat model, we asked them to list the possible consequences of ignoring each of the warnings. Responses that specifically mentioned fraud, identity theft, stolen credentials (or other personal information), phishing, or eavesdropping were coded as being correct. We coded as correct 39% of responses for FF2 warnings, 44% of responses for FF3 warnings, and 37% of responses for IE7 warnings.

Among FF2 and FF3 users, we observed no statistically significant differences based on whether they understood the consequences of one warning more than another. However, IE7 users were significantly more likely to misunderstand the consequences of ignoring the expired certificate warning than the other two warnings ($Q_2 = 6.500$, $p < 0.039$); 33% misunderstood the consequences of the expired certificate warning, whereas 40% misunderstood them for the unknown CA warning and 38% for the domain mismatch warning.

Incorrect responses fell into two categories: respondents who had no idea (or said there were no consequences) and respondents who mentioned other security threats. Many of those in the latter category mentioned viruses and worms. While it is possible that a malicious website may exploit software vulnerabilities or trick visitors into downloading malware, we considered these outside the scope of our survey because they usually either impact only users of a specific software version—in the case of a vulnerability—or they rely on the user taking additional actions—such as downloading and executing a file. Several responses mentioned malware but additionally claimed that those using up-to-date security software are not at risk. Others claimed they were not at risk due to their operating systems:

> "I use a Mac so nothing bad would happen."
> "Since I use FreeBSD, rather than Windows, not much [risk]."
> "On my Linux box, nothing significantly bad would happen."

[6] We discovered a typo in the survey: the FF2 unknown CA warning displayed an incorrect domain name towards the bottom of the message—the second time in the warning that the domain was displayed—which may have made this warning appear similar to a domain mismatch warning to those who read up to this point. However, based on the responses, only two of the 117 respondents (1.7%) noticed this error so we therefore concluded that it did not affect our results.

	Expired Certificate	Unknown CA	Domain Mismatch	
FF2	37%	45%	54%	$\chi^2_2 = 25.19 \quad p < 0.0005$
FF3	42%	52%	50%	$\chi^2_2 = 13.47 \quad p < 0.001$
IE7	47%	52%	53%	$\chi^2_2 = 12.79 \quad p < 0.002$

Table 5.2: Mean perceptions of the likelihood of "something bad happening" when ignoring each warning, using a 5-point Likert scale ranging from 0 to 100% chance. A Friedman test yielded significant differences for each browser.

	Expired Certificate	Unknown CA	Domain Mismatch	
FF2	1.70	2.10	2.29	$\chi^2_2 = 20.49 \quad p < 0.0005$
FF3	1.96	2.36	2.32	$\chi^2_2 = 9.00 \quad p < 0.011$
IE7	2.14	2.36	2.34	$\chi^2_2 = 16.90 \quad p < 0.0005$

Table 5.3: Mean perceptions of the consequences of ignoring each of the three warnings, using a 5-point Likert scale ranging from 0 to 4. A Friedman test shows that respondents in every web browser condition were likely to assign significantly lesser consequences to ignoring the expired certificate warning than when ignoring either of the other two warnings.

Of course, operating systems or the use of security software do not prevent a user from submitting form data to a fraudulent website, nor do they prevent eavesdropping. We further examined risk perceptions by asking participants to specify the likelihood of "something bad happening" when ignoring each of the three warnings, using a 5-point Likert scale ranging from "0% chance" to "100% chance." Comparing the responses for each of the warnings we found significant differences for all three web browsers: respondents consistently ranked the expired certificate warning as being less risky than both of the other warnings. Table 5.2 depicts the perceived likelihood of risk for each of the web browsers and each of the three SSL warnings.

We compared the perceived risk for ignoring the domain mismatch and unknown CA warnings and did not observe significant differences between FF3 and IE7, however we did observe that FF2 users perceived significantly more risk when viewing the domain mismatch warning.

To further examine whether there were differences in risk perception based on the destination website, we asked respondents to quantify the severity of the consequences of ignoring each of the SSL warnings using a 5-point Likert scale that ranged from "none" to "moderate" to "severe." As shown in Table 5.3, we found that respondents in every web browser condition were likely to assign significantly lesser consequences to ignoring the expired certificate warning than when ignoring either of the other two warnings.

The Role of Context

We expected respondents to have more reservations about ignoring a warning when visiting a website that was likely to collect personal or financial information than when visiting a website that was unlikely to collect such information. We examined the question of whether participants would be more likely to proceed past the warnings to *amazon.com* or *craigslist.org* by analyzing the data presented in Figure 5.1 based on which website was shown to participants. Using a chi-square test we found no significant differences between the two websites for any of the web browsers. Once we removed the responses coded as "maybe," so that we would be left with just the definitive answers, we discovered that when viewing the expired certificate warning while using FF3, respondents were significantly more likely to say they would discontinue visiting

	Tech score		Expired	Unknown CA		Domain Mismatch	
FF2	$\mu = 0.61$	*Experts*	69%	44%		31%	
	$\sigma = 1.14$	*Non-Experts*	63%	48%		31%	
FF3	$\mu = 0.99$	*Experts*	52%	13%	$\chi^2_2 = 12.37$	10%	$\chi^2_2 = 11.42$
	$\sigma = 1.42$	*Non-Experts*	47%	41%	$p < 0.002$	31%	$p < 0.003$
IE7	$\mu = 0.47$	*Experts*	42%	33%		29%	
	$\sigma = 1.02$	*Non-Experts*	36%	31%		29%	

Table 5.4: Percentage of experts and non-experts who said they would continue past the warnings. The first column shows respondents' average tech scores.

amazon.com than *craigslist.org* (32% said they would not proceed in light of the expired certificate on *craigslist.org*, whereas 53% said they would not proceed on *amazon.com*; $p < 0.028$ for a one-tailed Fisher's exact test). We believe that because respondents were not actually visiting websites and because they were personally not at risk, their stated behaviors may not actually mirror their observed behaviors.

The Role of Expertise

Finally, we wanted to examine whether respondents' level of technical expertise influenced their decisions to heed or ignore the warnings. As described in Section 5.1.1, we asked respondents a series of five questions to gauge their technical qualifications. We assigned each respondent a "tech score" corresponding to the number of questions they answered affirmatively. The first column of Table 5.4 lists the average scores for each of the web browser conditions.

For the purpose of comparing our "experts" with the rest of our sample, we classified as experts those who had scores of two or more. Those above the cut-off represented the top 16.7% of FF2 users, the top 26.5% of FF3 users, and the top 12.2% of IE7 users. We compared our "experts" to the rest of our sample (i.e. respondents with scores of zero or one) and found that responses did not significantly differ in most cases. We found significant differences only among FF3 users when viewing the unknown CA and domain mismatch warnings: experts were significantly less likely to proceed to the websites (Table 5.4).

Finally, we examined whether the experts were better able to identify the individual warnings than the rest of the sample. We found that while the experts were more likely to identify the warnings than non-experts, even in the best case, the experts were only able to correctly define the expired certificate warnings an average of 52% of the time, the unknown CA warnings 55% of the time, and the domain mismatch warnings 56% of the time. This indicates that either our metric for expertise needs to be improved, or that regardless of technical skills, many people are unable to distinguish between the various SSL warnings.

Conclusion

Our survey showed how risk perceptions are correlated with decisions to obey or ignore security warnings and demonstrated that those who understand security warnings perceive different levels of risk associated with each warning. These preliminary findings suggests that current SSL warnings suffer from problems at both the *Comprehension* and *Attitudes and Beliefs* stages of the C-HIP model because users may confuse one particular SSL warning with a different SSL warning. Likewise, these warnings may also suffer from problems at the *Motivation* stage of the C-HIP model when users do not understand the associated risks of

ignoring a warning. However, a limitation of surveys is they collect participants' self-reported data about what they think they would do in a hypothetical situation. Thus, it is useful to validate survey findings with experimental data.

5.2 Laboratory Experiment

5.2.1 Methodology

We conducted a laboratory study to determine the effect SSL warnings have on user behavior during real tasks. The study was designed as a between-subjects experiment with five conditions: FF2 (Figure 5.2(a)), FF3 (Figure 5.3), IE7 (Figure 5.2(b)), a single-page redesigned warning (Figure 5.4(b)), and a multi-page redesigned warning (Figure 5.4). Participants were asked to find information using four different types of information sources. Each task included a primary information source—a website—and an alternate source which was either an alternative website or a phone number. The primary information source for two of the tasks, the Carnegie Mellon University (CMU) online library catalog and an online banking application, were secured by SSL. We removed the certificate authorities verifying these websites from the trusted authorities list in each browser used in the study.[7] Therefore, participants were shown an invalid certificate warning when they navigated to the library and bank websites. We noted how users reacted to these warnings and whether they completed the task by continuing to use the website or by switching to the alternative information source. Finally, users were given an exit survey to gauge their understanding of and reaction to the warnings.

Recruitment

We recruited participants by posting our study on the experiment list of the Center for Behavioral Research at CMU. We also hung posters around the CMU campus. Participants were paid $10–20 for their participation.[8] All recruits were given an online screening survey (Appendix F), and only online banking customers of our chosen bank were allowed to participate. The survey included a range of demographic questions and questions about general Internet use.

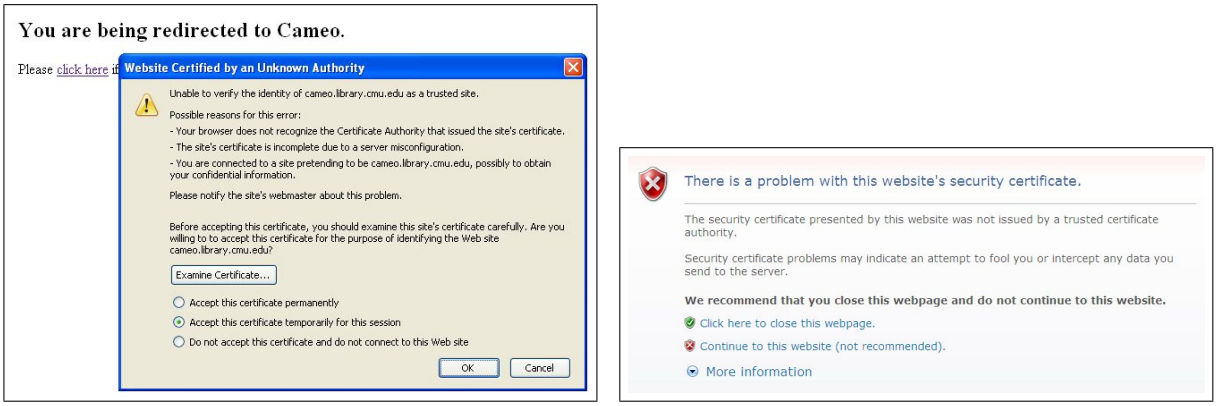
In total, 261 users completed our screening survey and 100 users qualified and showed up to participate in our study. We randomly assigned 20 users to each condition. Half the users in each condition were given the bank task first and half were given the library task first. Participants took 15–35 minutes to complete the study including the exit survey.

We tried to ensure that participants were not primed to think about security any more than they would in their natural environments. The study was presented not as a security study, but as a "usability of information sources study." Our recruitment postings solicited people who were "CMU faculty staff or students" and had "used online banking in the last year." However, we also required that participants have "purchased an item online in the last year" and "used a search engine" to avoid focusing potential participants on the banking tasks. Finally, our screening survey asked a series of questions, unrelated to the bank or CMU, whose responses were not used to screen participants.

[7] Ideally we would have performed a man-in-the-middle attack, for example by using a web proxy to remove the websites' legitimate certificates before they reached the browser. However, due to legal concerns, we instead simulated a man-in-the-middle attack by removing the root certificates from the web browser.

[8] Initially participants were paid $10, but we raised the payment to $20 to reach our recruiting goals.

Conditions



(a) Firefox 2 (b) Internet Explorer 7

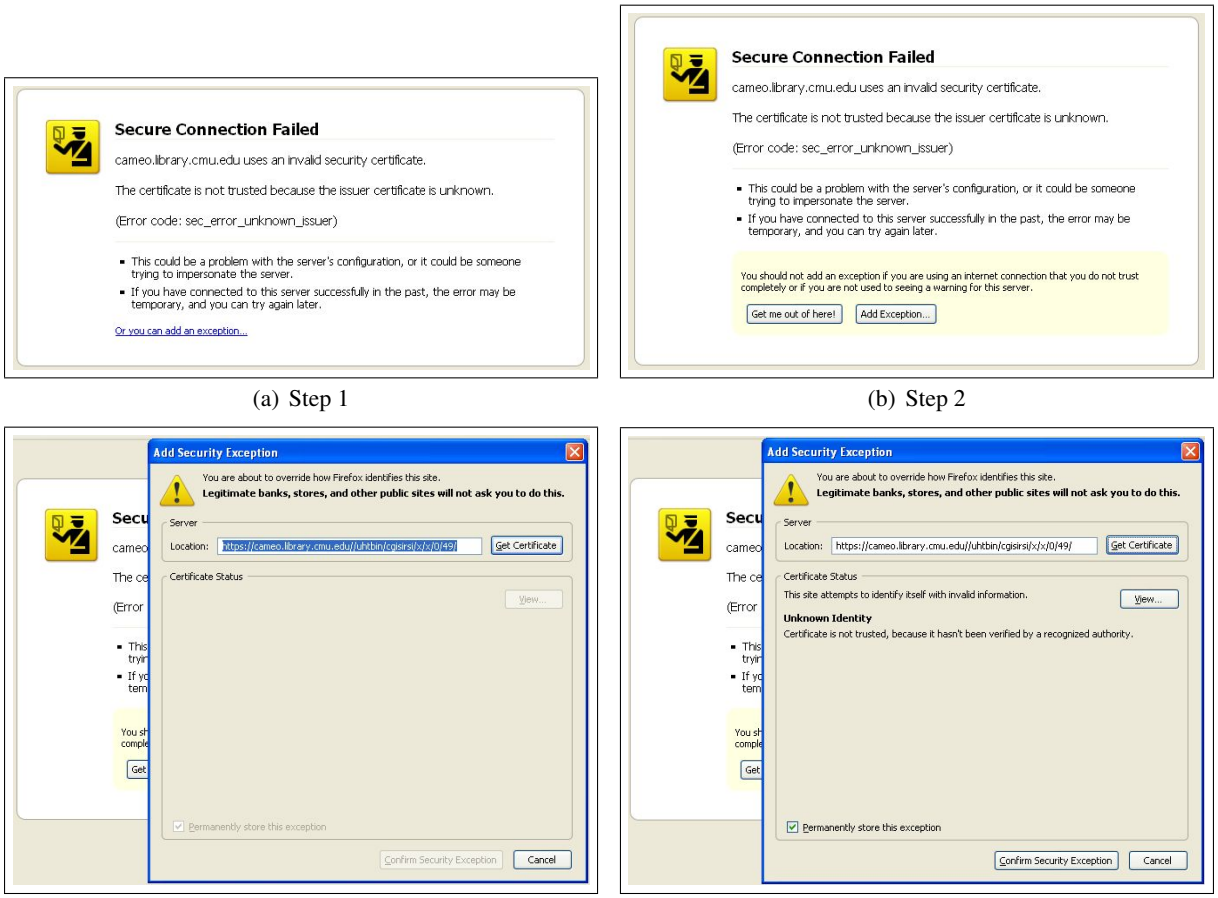
Figure 5.2: Screenshots of the FF2 and IE7 warnings.

The FF2 warning, displayed in Figure 5.2(a), is typical of invalid certificate warnings prior to 2006. This warning has a number of design flaws. The text contains jargon such as, "the site's certificate is incomplete due to a server misconfiguration." The look and feel of the warning, a grey dialog box with a set of radio buttons, is similar to a lot of other trivial dialogs that users typically ignore, such as "you are sending information unencrypted over the internet." The default selection is to accept the certificate temporarily. This is an unsafe default for many websites, including those of financial institutions like the online banking application in our study.

A more subtle problem with the FF2 warning, and those like it, is that it asks the user a question that they cannot answer. This problem is discussed at length by Firefox project co-founder Blake Ross in his essay entitled *Firefox and the Worry Free Web*. The warning asks the user to determine if the certificate problem is the result of a server/browser configuration problem or a legitimate security concern. Since users are not capable of making this determination, the dialog is "a dilemma to users." Ross calls on browser designers to do everything possible to make decisions for their users. When designers have to ask questions of their users, they should ask questions that users can answer [109].

The FF3 warning should be more noticeable to users than its predecessor because it takes over the entire page and forces users to make a decision. Additionally, it takes four steps to navigate past the warning to the page with the invalid certificate. First the user has to click a link, mysteriously labeled "or you can add an exception..." (Figure 5.3(a)), then click a button (Figure 5.3(b)), which opens a dialog requiring two more button clicks (Figures 5.3(c) and 5.3(d)). The first version of the FF3 warning required 11 steps [14]. This clearly represented a decision by Firefox developers that all invalid certificates are unsafe. They made the original version of the warning so difficult for users to override, that only an expert would be likely to figure out how to do it. While FF3 was in alpha and beta testing many users erroneously believed the browser was in error when they could not visit websites that they believed to be legitimate [15].

The IE7 warning, shown in Figure 5.2(b), occupies the middle ground between the FF2 and FF3 warnings. It takes over the entire page and has no default option, but differs from the FF3 warning because it can be overridden with a single click on a link labeled "Continue to this website." It has a slightly scarier look



(a) Step 1 (b) Step 2

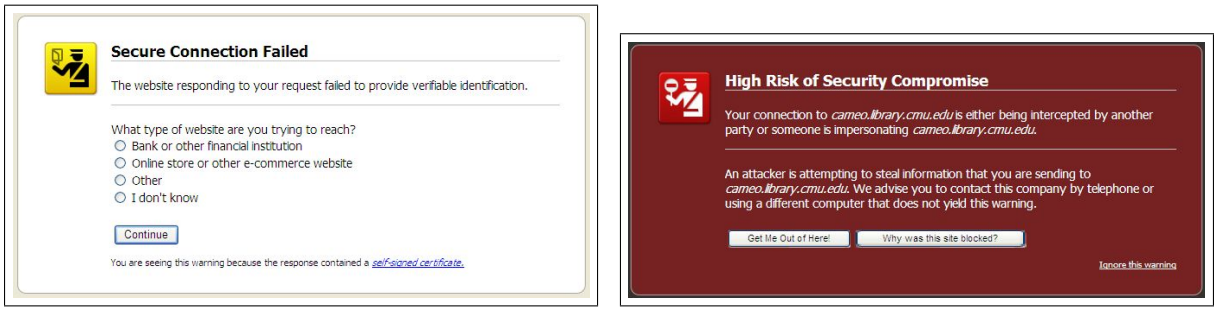
(c) Step 3 (d) Step 4

Figure 5.3: Screenshots of the four steps fo the FF3 warning.

and feel than the FF2 warning: the background color has a red tint and a large X in a red shield dominates the page. The warning also explicitly recommends against continuing. Finally, when viewing this warning the background of the address bar is red and continues to be red after one overrides the warning.

We designed two warnings using techniques from the warning literature and guided by results from our survey. Our multi-page warning first asks the user a question, displayed in Figure 5.4(a), and then, depending on the response, delivers the user either to the severe warning page shown in Figure 5.4(b) or to the requested website. The second version of the warning shows only the severe warning (Figure 5.4(b)). Both versions were implemented in IE7. We used the resourcemodify tool [107] to replace the HTML file of the native warning in an IE dll with our HTML files.

The second version of our warning serves two purposes. First, it attempts to see how users react to a simple, clear, but scary warning. The warning borrows its look and feel from the FF3 phishing warning. It is red and contains the most severe version of Larry [131] the Firefox "passport officer." The title of the page is clear and harsh: "High Risk of Security Compromise." The other context is similarly blunt (e.g. "an attacker is attempting to steal information that you are sending to *domain name*."). Even the default button,



(a) Page 1 (b) Page 2

Figure 5.4: Screenshot of redesigned warning.

labeled "Get me out of here!" signifies danger. The only way for a user to continue is to click the tiny link labeled "Ignore this warning" in the bottom right corner. The second purpose of the single page warning is to help us interpret the results from our multi-page warning. We compare the multi-page results to the single-page results to see how the question affects user actions independent of the the scary second page.

The original FF3 warning aimed to avoid asking users any question, and instead decide on users' behalf that invalid certificates are unsafe. However, even the Firefox designers eventually realized this could not work in the real world because too many legitimate sites use self-signed certificates. Instead, our warning aims to ask the users a question that they can answer and which will allow us to recommend an action to them. Our question is, "What type of website are you trying to reach?" Users were required to select from one of four responses: bank or other financial institution, online store or other e-commerce site, other, and I don't know. If the user selects bank or online store they see the severe warning that discourages them from continuing. If they select other or I don't know they proceed immediately to the website.

We chose to only show the second warning page for financial institutions and online stores. Many other sites use certificates signed by well-known certificate authorities (e.g. VeriSign, Thawte). One might think our warning makes such sites less safe. However, banks and online stores are the most attacked websites [104], so protecting these sites will have the most impact. In addition, participants in our pilot studies ignored almost all existing warnings. Therefore, we hypothesized that our warning would be no worse than existing warnings at protecting users against self-signed certificate attacks against sensitive sites that are not banks or e-commerce sites.

Experimental Setup

All studies were conducted in our laboratory on the same model of laptop. Each participant's interaction with the laptop took place within a virtual machine which was reset to a snapshot after study completion. This ensured that all browser and operating system settings were exactly the same for every participant and that any sensitive data entered by the participant and stored on the machine (e.g. bank password) was destroyed after the participant left. All instructions were read from the exact same script. Finally, participants were told that "we want to see how people interact with information sources on their own, so the experimenter will not be able to help you figure out how to complete a task."

Tasks

After participants signed IRB consent forms, the experimenter handed them an instruction sheet and read this sheet aloud. Participants were reminded that they would be "visiting real websites and calling real organizations" and therefore should go about "each task in the way you would if you were completing it with the computer you usually use." Participants were also instructed to "think aloud and tell us what you are thinking and doing as you complete each task," in order to give us qualitative reactions to the warnings. The experimenter took notes throughout the study. The study was recorded (audio only), which allowed experimenters to retrieve details that were missed during note taking.

After the instructions were read and digested, the instruction sheets for each task were handed to the participant and read aloud by the experimenter one by one. The tasks were not revealed before the study, nor was the next task revealed until all previous tasks had been completed. The first task asked participants to find the total area of Italy in square kilometers using Google or Ask.com as an alternative. The second task was to look up the last two digits of the participant's bank account balance using the online banking application or using phone banking. The third task was to locate the price of the hardcover edition of the book *Freakonomics* using Amazon.com or the Barnes and Noble website. Finally, the fourth task was to use the CMU online library catalog or alternatively the library phone number to retrieve the call number of the book *Richistan*.

The first and third tasks were "dummy tasks," since the bookstore and search engine revealed no warnings. Instead, they reinforced to participants that the goal of the study was information sources, not security. Half the participants in each condition had the second and fourth tasks—the warning tasks—swapped so that we could control for the ordering of the warnings.

Researchers have found that user study participants are highly motivated to complete assigned tasks. Participants want to please the experimenter and do not want to "fail" so they sometimes exert extreme effort to complete the task [90, 68]. A study closely related to ours, the *Emperor's New Security Indicators* [110], was criticized for not taking into account this "task focus" phenomenon [102]. Critics worried that participants were ignoring the warnings in the study because of task focus and not because this is what they would do in a more natural environment.

Our study design mitigates participants' task focus by presenting an alternate method for each task so that participants could "pass the test" without ignoring the warnings. We instructed participants to "try the suggested information source first," to ensure that participants would only call the library or bank as a reaction to the warning. As there were no obstacles to completing the dummy tasks using the suggested information source, none of the participants used the alternate method to perform the dummy tasks.

Exit Survey

After completing all four study tasks, participants were directed to an online exit survey hosted by SurveyMonkey (Appendix G). In the exit survey we asked 45 questions in six categories. The first set of questions asked about their understanding of and reaction to the bank warning in the study. The second set of questions asked the same questions about the library warning. The third set asked questions to gauge their general understanding of certificates and invalid certificate warnings. The fourth set gauged participants prior exposure to identity theft and other cyberthreats. The fifth set, which were also asked in the online SSL survey, asked them about their technical experience including their experience with computer security. Finally, the sixth set asked general demographic questions like age, gender and education level.

	FF2		FF3		IE7		Single-Page		Multi-Page	
Bank	18	(90%)	11	(55%)	18	(90%)	9	(45%)	12	(60%)
Library	19	(95%)	12	(60%)	20	(100%)	16	(80%)	19	(95%)

Table 5.5: Number (and percentage) of participants in each condition who ignored the warning and used the website to complete the library and bank tasks.

5.2.2 Results and Analysis

The primary goal of any SSL warning should be to prevent users from transmitting sensitive information to suspicious websites. A secondary—but still important—goal is to allow users to continue in the event of a false positive (i.e. when a certificate error is unlikely to result in a security compromise). In our study we examined these goals by observing whether participants discontinued visiting the bank website while continuing to the library website. These results from our laboratory experiment are displayed in Table 5.5. Participants who saw our single-page or multi-page warnings were more likely to heed the warnings than participants who saw the FF2 or IE7 warnings, but not the FF3 warning. In contrast, participants who saw our multi-page warning were more likely to visit the library website than participants who saw the FF3 warning. In the rest of this section we discuss demographics, present more detailed comparisons of the conditions and tasks, and present interesting qualitative results from our exit survey.

Participant Characteristics

We did not find any statistically significant demographic imbalances between participants in our randomly assigned conditions. The factors we tested were gender, nationality, age, technical sophistication, and a metric we call "cyberthreat exposure" designed to measure participants' prior experiences with information theft and fraud. Most demographic factors were determined by a single exit survey question (e.g. gender, nationality). Technical sophistication was measured by a composite score of five question, the same as in the online survey. Similarly, cyberthreat exposure was measured by asking participants if they have ever had any account information stolen, found fraudulent transactions on bank statements, had a social security number stolen, or if they had ever been notified that personal information had been stolen or compromised.

Our participants were technically sophisticated, mostly male, and mostly foreign students. We had 68 male and only 32 female participants. All of our participants were between the ages of 18–30, and all but two were students. Sixty-nine participants were born in India, 17 in the United States, and the remaining were from Asia (10) and Europe (4). The average tech score was 1.90, which is significantly larger than the 0.66 average among the survey respondents.

We do not have a large enough sample size to determine whether age, profession, or nationality influenced participant behavior. In addition, our participants had so little cyberthreat exposure—83 participants answered affirmatively to 0 out of 4 questions—that we could not determine if exposure correlated with our results. On the other hand, while our sample was large enough to observe behavioral differences based on gender and technical sophistication if large differences existed, we observed no statistical differences in participant behavior based on those factors. Finally, we found no statistical difference in behavior based on task order in any of the conditions.

Effect of Warning Design on Behavior

Our study design focused on evaluating whether SSL warnings effectively prevent users from transmitting sensitive information to suspicious websites, while allowing them to continue in the event of a false positive.

We hypothesized that participants visiting the bank website who see our redesigned warnings would be significantly more likely to discontinue than participants who see the other warnings. We used a one-tailed Fisher's exact test to analyze our results. We found that significantly more participants obeyed our single page warning than obeyed the FF2 and IE7 warnings ($p < 0.0029$ for both comparisons). Similarly, our multi-page warning performed better than the FF2 and IE7 warnings ($p < 0.0324$). However, FF3 was equivalently preventative, and it was in fact significantly better than the FF2 and IE7 warnings ($p < 0.0155$).

We also hypothesized that participants visiting the library website who see our redesigned warning will be significantly more likely to continue than participants who see the other warnings. In this case our hypothesis turned out to be mostly false. Participants who viewed our multi-page warning were significantly more likely to use the library website than participants who saw the FF3 warning ($p < 0.0098$). However, our multi-page warning was not better than the FF2 or IE7 warnings and our single page warning was better than none of the other warnings. The FF3 warning caused significantly more participants to call the library than the FF2 warning ($p < 0.0098$) or the IE7 warning ($p < 0.0016$).

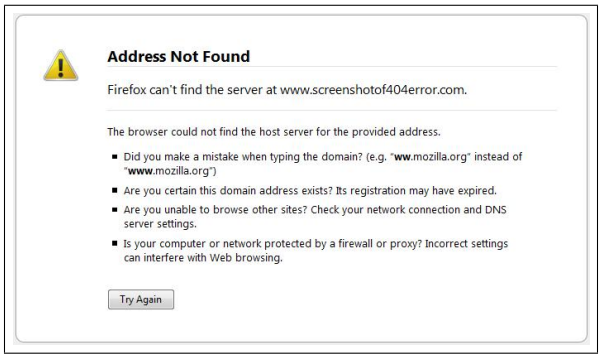


Figure 5.5: Screenshot of server not found error in FF3.

Two participants in the FF3 condition and one in our multi-page warning condition thought the library and bank servers were down or that we had blocked their websites. One wrote in the exit survey "the graphics made me feel the server was down" and another wrote "I just saw the title and assumed that it is just not working on this computer." We suspect that users confuse the warnings with a 404 or server not found error, like the one shown in Figure 5.5. The warnings have very similar layouts and coloring. The yellow Larry icon in the FF3 warning (Figure 5.3(a)) and the first page of our multi-page (Figure 5.4(a)) warning is similar to the yellow triangle in Figure 5.5.

We took careful note of how participants in the multi-page warning condition answered the question "What type of website are you trying to visit?" presented to them on the first page of the warning. Fifteen participants answered exactly as expected – they selected "other" for the library and "bank or other financial institution" for the bank. The remaining five participants exhibited noteworthy behaviors: one participant did not answer the question for either task, while three participants performed the library task first and appropriately answered "other," but also inaccurately answered "other" when visiting the bank website. This is stark evidence of the ill-effects of warning habituation – these participants learned how to ignore the warn-

Condition	Read		Didn't Read		Understood		Didn't Understand	
	Logged In	Called	Logged In	Called	Logged In	Called	Logged In	Called
FF2	4	2	14	0	7	2	11	0
FF3	2	2	9	7	4	2	7	7
IE7	4	1	14	1	8	2	10	0
Single-Page	4	6	5	5	4	7	5	4
Multi-Page	8	6	4	2	7	6	5	2

Table 5.6: Behavior in the bank task by reading, understanding, and condition.

ing in the library task and immediately reapplied their knowledge to the bank task. Finally, one participant first performed the bank task and correctly answered "bank or other financial institution." However, when she saw the second page of the warning she clicked the back button and changed her answer to "other."

Risk Perception in Context

We hypothesized that participants who viewed our multi-page warning would be more likely to obey the warnings when they were visiting the bank website than when they were visiting the library website. Because this warning took context into account in determining severity, it appeared to be more severe on the bank website. All 14 participants in our study who heeded the library warning also heeded the warning at the bank. An additional 18 participants heeded the bank warning and proceeded past the library warning. Participants who viewed our multi-page warning ($p < 0.0098$) and our single-page warning ($p < 0.0242$) were significantly more likely to heed the warning at the bank than at the library.

We believe the behavior exhibited by users of our single page warning can be explained both by its success in raising awareness of risk and its clear communication of what users should do in response to the risk. When the 11 participants who heeded the single-page bank warning were asked in the exit survey "Why did you choose to heed or ignore the warning?" 9 out of 11 specifically mentioned the security of their information as the reason. In contrast only 2 participants in each of the FF2, FF3, and IE7 conditions mentioned risk in response to the same question. In addition, 10 of the 20 participants in the our single-page warning condition when asked, "What action(s) did you think the warning at the bank wanted you to take?" responded that it wanted them *not* to proceed. Only 3 FF2, 2 FF3, and 4 IE7 participants answered the same way.

Impact of Reading and Understanding

In each of the first two sections of the exit survey we asked participants if they "read the text of the warning at the *bank/library* site." At the bank website, significantly more people read our multi-page warning than the FF2 ($p < 0.0128$), FF3 ($p < 0.0018$), or IE7 ($p < 0.0052$) warnings (Table 5.6). There were no other significant differences in reported readership across conditions or tasks. We used a chi-square test to see if there was a difference in how reading affected behavior. Among the participants who did not read the warnings, FF2 and IE7 users were significantly more likely to log in to the bank website ($\chi^2_4 = 13.56$, $p < 0.009$), whereas FF3 users were significantly less likely to log in to the library website ($\chi^2_4 = 18.38$, $p < 0.001$).

The exit survey asked participants "what did you believe the warning at the *bank/library* website meant?" Answers were entered into a free response text box and we categorized the responses according to whether or not they demonstrated understanding of the warning, as we had done in the survey (Table 5.6). In particular, participants who wrote that their connection may be compromised or that the identity of the destination website could not be verified were deemed to understand the warning. All other responses were coded as not understanding the meaning. There were no significant differences in the number of participants who understood the warnings based on condition in either task. However, participants in the FF3 condition who did not understand the warning were significantly more likely to call than users in the FF2 ($p < 0.0078$) and IE7 ($p < 0.0188$) conditions. Seven of the 14 participants who did not understand the FF3 warning called the bank. This is evidence that the FF3 users may have been prevented from visiting the websites because they did not know how to override warnings, and not because they understood the risks of proceeding.

One expects that participants who claimed to have read the warnings would be more likely to understand their meeting. When we combined the data from just our two warnings, single-page and multi-page, we found a statistically significant correlation ($p < 0.020$). However, we do not have enough data to determine whether there is a correlation for the three native warnings (FF2, FF3, and IE7).

Other Observations

One worry for browser designers trying to design effective warnings is that they will cause users to switch browsers, in favor of a browser that shows a less severe warning. In fact, during our study a few participants who viewed our warnings or the FF3 warnings asked or attempted to perform one of the tasks in a different browser. We directed them to continue using the browser they had been using. No participants in the FF2 and IE7 conditions tried to switch browsers. This indicates that complex warning designs may cause a small number of users to switch browsers. Therefore, for the sake of these users' security, it may be best if all browsers converged on a single warning design.

Response	FF2	FF3	IE7	Single-Page	Multi-Page	Total
Yes	8	7	10	4	1	30
No	8	11	5	16	16	56
Not sure	4	2	5	0	3	14

Table 5.7: Number of participants in each condition who claimed to have seen the warning before at the bank.

Among our strangest results were the answers to the questions: "Before this study, had you ever seen the warning you saw at the *bank/library* web site?" (Table 5.7). A total of 30 participants said they had seen the warning before at the bank website compared to only 16 at the library website. Of these, five participants in the bank task thought they had seen our warnings before. We do not think 30% of our participants have been scammed by man-in-the-middle attacks at their bank and we know for sure that the 5 participants had never seen our warnings before. This is dramatic evidence of memory problems, warning confusion, and general confusion with regard to certificate errors.

In the exit survey we asked participants to use a 7-point Likert scale to report the influence of several factors on their decision to heed or ignore the warnings. The factors we included were: the text of the warning, the colors of the warning, the choices that the warning presented, the destination URL, and the look and feel of the destination website. We expected significantly more participants to grade the color and

text of the website highly for our warnings. However, there was no statistically significant difference in participants' responses based on condition.

	FF2	FF3	IE7	Single-Page	Multi-Page
Avg. # of Hesitation Actions	0.1	1.8	0.4	1.2	1.4

Table 5.8: Hesitation actions by condition.

For each participant, the experimenters noted the "hesitation actions" performed by each participant. These actions included retyping the URL, refreshing the page, searching for the website (i.e. querying a search engine), clicking the back button, and selecting a help link or button. As we expected, the complex and unfamiliar warnings—our warnings and the FF3 warning—yielded many more hesitation actions than the simple IE7 and FF2 warnings (see Table 5.8).

5.3 Discussion

Our warnings markedly improved user behavior, but all warning strategies, including ours, leave too many users vulnerable to man-in-the-middle attacks. The five warnings we evaluated embodied three different strategies: explain the potential danger facing users, make it difficult for users to ignore the warning, and ask a question users can answer. The strategies have differences that we will discuss later in this section. However, regardless of how compelling or difficult to ignore, users think SSL warnings are of little consequence because they see them at legitimate sites. Many users have a completely backward understanding of the risk of man-in-the-middle attacks and assume that they are *less* likely to occur at trusted sites like those belonging to banks. If they do become fraud victims, they are unlikely to pinpoint it to their decision to ignore the warning, and therefore are unlikely to learn from their mistakes. Thus, user's attitudes and beliefs about SSL warnings are likely to undermine their effectiveness [32]. Therefore, the best avenue we have for keeping users safe may be to avoid SSL warnings altogether and *really* make decisions for users—blocking them from unsafe situations and remaining silent in safe situations.

5.3.1 Explain the Danger

The FF2, IE7, and our single page warning take the standard tactic of explaining the potential danger to users. The FF2 warning, which is an unalarming popup box with obscure language prevented very few users from visiting the bank or library. The IE7 warning which has clearer language and a more frightening overall look did not perform any better. On the other hand, our single page warning, with its black and red colors was the most effective of the five warnings at preventing users from visiting the bank website. The red and black design caught users' attentions—preventing errors at the *Attention Switch and Maintenance* stages of the C-HIP model—and also stood out from less serious warnings, preventing errors at the *Attitudes and Beliefs* stage of the C-HIP model. Our single page warning also helped users to understand the risks of continuing to the bank website by stating them clearly, thereby minimizing errors at the *Motivation* stage of the C-HIP model [140]. At the same time, only four users called the library, indicating that our single-page warning would be only a minor nuisance for legitimate websites. That said, without conducting further studies we do not know if our single page warning will become less effective as users are habituated to it when visiting legitimate websites.

5.3.2 Make it Difficult

The FF3 warning, as discussed at length in Section 5.2.2, prevents user from visiting websites with invalid certificates by confusing users and making it difficult for them to ignore the warning. This improves user behavior in risky situations like the bank task, but it presents a significant nuisance in legitimate situations like the library task. Many legitimate websites that use self-signed certificates have posted online tutorials teaching users how to override the FF3 warning (see e.g. [70, 156, 26]). We suspect that users who learn to use the warning from these tutorials, by simple trial and error, help from a friend, etc., will ignore subsequent warnings and will be left both annoyed and unprotected. Thus, these warnings only protect users because they are difficult to understand, and not because they understand the risks of proceeding.

5.3.3 Ask a Question

Our multi-page warning, introduced in Section 5.2.1, asks the user a question. Fifteen of the 20 users answered correctly at the bank website and 20 of 20 answered correctly at the library site. As discussed in Section 5.2.2, we believe the few that did not, either knowingly gave the wrong answer to reach the destination website or confused the warning with a server unavailable error. This indicates that our question is something users can easily answer. Our multi-page warning was not better than the FF3 or our single-page warning at the bank site. However, it was significantly better than the FF3 warning at the library site and slightly better than the single-page warning. More importantly, users will only see our warning at bank and e-commerce sites extremely rarely so they will not become habituated to the scary second page of the warning. We hope that this will help the warning maintain its effectiveness.

5.3.4 Avoid Warnings

The ideal solution to SSL warning problems is to block access when users are in true danger and allow users to proceed when they are not. This ideal is probably unattainable, but two systems recently presented by the research community, ForceHTTPS [75] and Perspectives [133], are steps in the right direction. ForceHTTPS allows website owners to specify whether browsers should terminate the connection on a given SSL error, thus taking the decision away from individual users and putting it in the hands of the individual website operators. For instance, it is likely a good practice to block users from accessing their bank websites when the connection yields errors. Perspectives uses “notaries” to continuously examine website certificates to determine whether a certificate has changed over time. Thus providing key continuity management—alerting users only when keys change [62, 67]. Both systems identify websites that are likely to be unsafe and use warnings to stop users from proceeding. It would be better to block these unsafe sites entirely. We expect both systems to have a extremely low false positive rates, therefore minimizing the number of security warnings that are shown to users, and therefore minimizing habituation effects.

Chapter 6

Privacy Information Timing Study

This chapter is largely a reproduction of a paper co-authored with Janice Tsai, Lorrie Faith Cranor, and Alessandro Acquisti [49]. Thanks to Daniel Rhim for his assistance carrying out this study. We are also grateful to the companies who participated: EdenFantasys.com, Instawares, Little Office Supply, NiteTimeToys.com, Office Quarters, On Time Supplies, SheVibe, and The Dirty Bunny. This work was supported in part by the National Science Foundation under grant CCF-0524189 and by U.S. Army Research Office contract no. DAAD19-02-1-0389 (Perpetually Available and Secure Information Systems) to Carnegie Mellon Universitys CyLab.

Previous studies have shown that users may be willing to pay a premium to know when they are visiting a high privacy website [129]. But there is still an open question of *how* to effectively convey website privacy information. We performed a laboratory study where we tightly controlled the price of two items offered by several online vendors such that participants would have to pay more money to purchase the items from vendors with better privacy policies—a privacy premium. We selected the privacy premiums based on the results of an online survey designed to determine the maximum amount online shoppers would be willing to pay for increased privacy. A total of 89 participants came to our laboratory and purchased two items using their own credit cards and providing their personal billing information. One item, a vibrating sex you, was selected to elicit heightened privacy concerns. The other item, a pack of Duracell AA batteries, was selected to elicit minimal privacy concerns. We created four conditions that used different privacy icons to annotate the websites such that we varied both *when* and *how* the icons were displayed.

Our results demonstrate, first, that many online shoppers will go to extra efforts to purchase from high privacy websites when privacy indicators are available. Second, we show that online shoppers who are less privacy-motivated will pay significantly more for privacy when privacy indicators are presented to them before visiting websites, rather than after they arrive at a website. Third, we demonstrate that online shoppers are more likely to take privacy indicators into account when purchasing privacy-sensitive items.

6.1 Privacy Premium Survey

Before our experiment, we conducted an online survey to estimate the maximum premium people would be willing to pay to purchase from a website with a high privacy level (Appendix H). We recruited 676 Internet

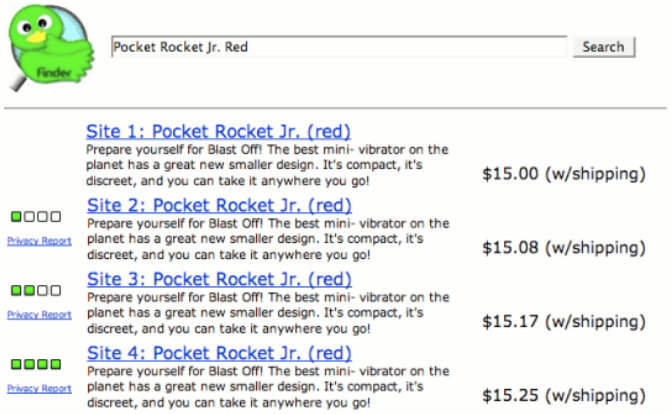


Figure 6.1: Example screenshot used in the privacy premium survey.

Indicator	Premium 1	Premium 2	Premium 3
☐☐☐☐	$15.00	$15.00	$15.00
■☐☐☐	$15.08	$15.25	$15.50
■■☐☐	$15.17	$15.50	$16.00
■■■■	$15.25	$15.75	$16.50

Table 6.1: The privacy premiums and associated privacy indicators used in the survey. The privacy indicator for the cheapest website was only displayed to half of the respondents.

users through Craigslist and sweepstakes websites in June 2008. As an incentive to participate, we offered participants a chance to win a $75 Amazon gift card. The survey contained five pages of Privacy Finder screenshots (Figure 6.1). Each screenshot depicted four search results for identical products with identical descriptions. The search results only differed based on the privacy indicator placed to their left and the price information placed to their right. Both the price and privacy level increased with each subsequent search result. Thus, the websites with the highest privacy ratings also had the highest prices. We assigned half the respondents to a between-group condition in which the cheapest website had no privacy indicator and the other half to a condition in which the cheapest website had the lowest privacy level. The product displayed in the search results alternated between the sex toy and pack of batteries that laboratory participants would be purchasing, with the order randomly selected. Respondents were given the following instructions:

> "Given only the information displayed in the search results, from which web site would you be most likely to make this purchase? Please choose one site, even if this is not a product you would be likely to ever purchase."

Respondents were exposed to two of three possible premiums for the highest privacy—denoted by four green boxes: $0.25, $0.75, and $1.50. The premiums and associated privacy indicators are shown in Table 6.1. The privacy premiums were randomly assigned so that respondents saw the same premium for the

first two pages (i.e. respondents saw the same premium for both products). The third page of the survey contained a control where one of the two products was randomly displayed with identical prices for each of the four search results. The privacy indicators varied so that we could examine whether participants would select the website with the highest privacy level in the absence of a premium.

The fourth and fifth pages followed the same protocol as the first and second pages, but participants were randomly assigned one of the two privacy premiums they had not already seen. However, we decided not to include these results in the analysis since we found evidence that participants' willingness to pay the subsequent premiums was highly dependent on the first premium to which they were exposed.

We combined the two between-group conditions for the analysis when we discovered that the only difference occurred when respondents encountered the highest privacy premium: those selecting the batteries were significantly more likely to select the first website—the cheapest one—when the indicator was absent ($t_{239} = 2.175, p < 0.031$).[1]

The ideal privacy premium for our laboratory study is the highest one that survey respondents would be willing to pay for both products; the survey responses likely provided an upper bound because the respondents reported how much they *would* pay without actually having to pay that amount. Using ANOVA to compare the three privacy premiums for each of the two products we found no significant differences between the three premiums when respondents considered the sex toy: most respondents indicated they were willing to pay any privacy premium presented to them. However, when the privacy premium was \$1.50, respondents were more likely to purchase the batteries from cheaper vendors, and therefore unwilling to pay a premium for privacy ($F_{2,673} = 6.251, p < 0.002$). At the same time, respondents indicated they were still willing to spend \$0.25 and \$0.75 for increased privacy when purchasing the batteries. We concluded a privacy premium of \$1.50 may be too high for our laboratory experiment.

A pairwise t-test confirmed that a \$0.75 privacy premium would still allow us to observe differences between the two products. Respondents indicated they were willing to spend significantly more money for the sex toy—in exchange for greater privacy—than for the batteries ($t_{214} = 5.226, p < 0.0005$). We concluded that a \$0.75 privacy premium would be low enough that laboratory participants would consider paying it for both products, while still allowing us to observe differences in behavior between the two product purchases.

6.2 Methodology

Our primary goal for this study was to examine whether the placement and timing of privacy indicators impacts purchasing decisions. In order to quantify differences in purchasing behaviors, we created a controlled privacy premium: participants who wanted a higher degree of privacy would have to pay a fixed amount for it. We also wanted to determine whether participants' behaviors would differ when purchasing a product that did not raise additional privacy concerns compared to a product that did. We designed the laboratory experiment to test the following hypotheses:

1. Participants will pay for increased privacy when they see privacy indicators.

[1] For a privacy premium of \$1.50, users may purchase from a website with an unknown privacy policy (i.e. the cheapest website) if the item being purchased does not raise privacy concerns.

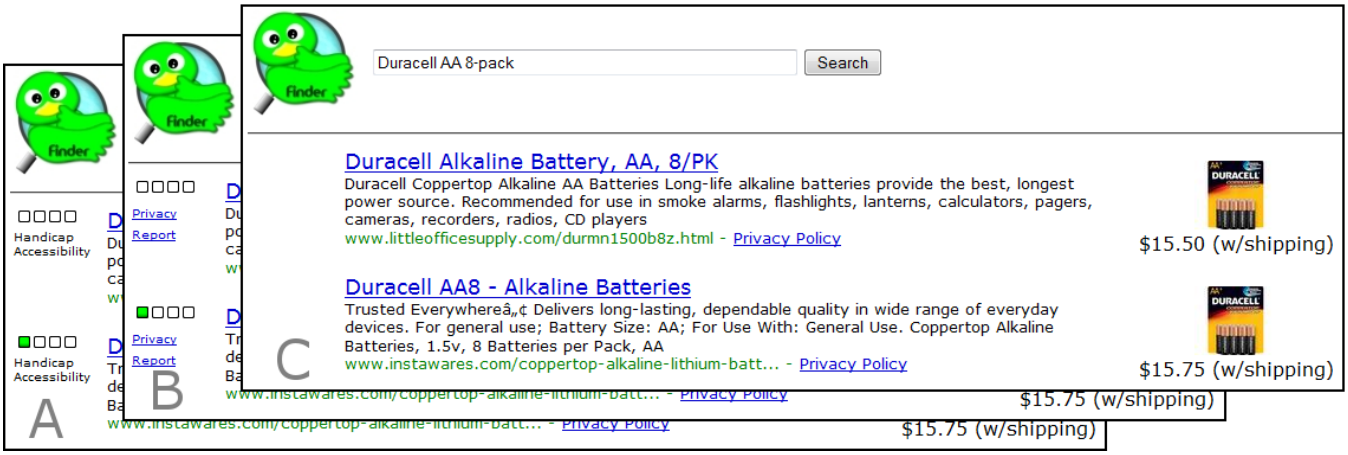


Figure 6.2: Screenshot of the search results for the four study conditions: (A) participants in the *handicap* condition saw the handicap accessibility indicators; (B) participants in the *privacy* condition saw the privacy indicators; and (C) participants in the *frame* and *interstitial* conditions did not have annotated search results.

2. Participants who see privacy indicators will pay more for the privacy-sensitive item than the item that does not raise additional privacy concerns.

3. Participants will be more likely to pay for increased privacy when they see privacy indicators alongside search results before visiting a website than when they see privacy indicators after clicking on search result links.

4. Participants will be more likely to pay for increased privacy when they see privacy indicators before they see the content of a website than when they see privacy indicators alongside the content of a website.

5. Participants who see privacy indicators after clicking on search result links will visit more websites than those who see privacy indicators alongside search results.

6.2.1 Study Design

We conducted a laboratory experiment during the summer of 2008 using participants from the Pittsburgh area. We recruited 89 participants using Craigslist and flyers on bus stops, telephone poles, and community bulletin boards. We used a screening survey to gather basic demographic data and to assess privacy concerns related to using the Internet and online shopping (Appendix I). Because the privacy indicators we tested were designed for use by individuals who have privacy concerns when shopping online, we used the same screening survey and screening methodology used in our previous study to screen out those who perceived little or no privacy risk when shopping online [129]. Based on this requirement, we screened out 16.39% of the responses (50 of 305).

We chose a specific vibrating sex toy, the "Pocket Rocket Jr.," as the privacy-sensitive item. We instructed participants to purchase the red version so that our results would not be confounded by the availability of differing colors from different vendors. We chose an 8-pack of Duracell AA batteries as the item

Hit #	Indicator	Price
1	☐☐☐☐	$15.50
2	■☐☐☐	$15.75
3	■■☐☐	$16.00
4	■■■■	$16.25
5		$16.75+

Table 6.2: The prices and privacy ratings for both sets of search results, the batteries and the sex toy. Participants who wanted the highest level of privacy had to pay an additional $0.75 for each product.

unlikely to raise additional privacy concerns beyond the act of providing personal information to an online vendor. We tightly controlled the price of each item by collaborating with four office supply vendors and four sex toy vendors who had varying privacy policies.[2] We asked the vendors to set specific prices based on their privacy policies and the results of our privacy premium survey.[3] Privacy Finder returned static results pages when specific search strings (or variants thereof) were submitted: "Pocket Rocket Jr. Red" and "Duracell AA 8-pack." Each of these two pages of search results contained five hits with varying prices and privacy ratings, as seen in Table 6.2. In both sets of search results we also included a fifth search result that did not have a privacy rating. This website had the highest price of the five and was included because we were curious if any participants would pay more than the $0.75 privacy premium to buy from a website with an unknown privacy policy, and whether they would understand that the lack of any indicator corresponds to an unknown privacy policy.[4]

We randomly assigned participants to one of three experimental conditions or the control condition, balancing the gender of participants in each condition:

- Handicap Accessibility (control): Participants were shown annotated search results (Figure 6.2A). However, we labeled the privacy indicators as "handicap accessibility" so that the indicators were not associated with privacy. The links to the privacy reports (i.e. the machine-generated privacy policy summaries) were removed.[5] We used this condition to examine whether participants in the other conditions were genuinely thinking about privacy or whether they were choosing websites simply based on the presence of irrelevant green indicators.

- Privacy (experimental): Participants were shown annotated search results with privacy indicators (Figure 6.2B).

- Frame (experimental): Participants were shown search results that were not annotated (Figure 6.2C).

[2]We contacted over twenty vendors for each product until four vendors for each product agreed to participate. For the vendors who lowered their prices, we compensated them for the difference. We only contacted vendors who participants were likely unfamiliar with; a full list of the vendors appears in the Acknowledgements.

[3]We used a privacy premium of $0.75 based on the results of the survey. Due to vendor constraints we had to set the base price at $15.50 rather than the $15.00 we used in the premium survey.

[4]No subject purchased either product from this website, and we therefore do not mention it in the analysis.

[5]Privacy reports are not discussed anywhere else in this paper since too few participants clicked them for us to draw any conclusions.

Figure 6.3: Screenshot of a website in the *frame* condition.

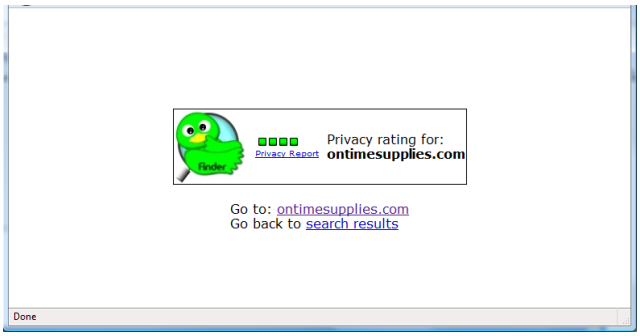


Figure 6.4: Screenshot of a website in the *interstitial* condition.

Once a participant visited a website from the search results, a frame appeared at the top of the website that displayed the privacy indicator and a link to the privacy report (Figure 6.3). We created this condition to simulate the Privacy Bird experience: users who wanted to comparison shop based on privacy indicators would have to visit a website in order to see its privacy rating. We hypothesized that users would find this tedious and therefore make poor privacy choices, especially when purchasing the batteries since they would likely be less motivated to protect their privacy.

- Interstitial (experimental): Participants were shown search results that were not annotated (Figure 6.2C). Once a participant visited a website from the search results, they saw an interstitial—a full screen message—with the privacy indicator (Figure 6.4). We created the interstitial condition to examine whether the content of a website detracted from the privacy indicator. We wanted to control for users being able to view website content alongside the privacy indicator in the *frame* condition. We hypothesized that users would choose higher privacy in this condition because they would be making the decision solely based on the privacy indicator.

We found no significant differences between the average ages ($\mu = 30.24, \sigma = 12.253$) of the groups. Differences paid for each product by gender were not significant ($t_{87} = 1.73$, $p < 0.087$ for the sex toy; $t_{87} = 0.96$, $p < 0.34$ for the batteries). We therefore believe the groups consisted of comparable populations.

Our flyers solicited participants for a study on the usability of an online search engine so that we would

not prime participants to privacy. The flyers informed participants that we would be paying them to shop online and that they would "Keep the Change!" When participants arrived for the experiment, we handed them instruction sheets that labeled the various features of Privacy Finder: the search box, the list of results, the annotated price information, the product pictures, and the privacy indicators. All references to "Privacy Finder" were changed to "Finder" in order to reduce priming effects. Likewise, we scheduled all participants at least 72 hours after taking our privacy concerns screening survey.

We gave participants packets that instructed them to complete several information retrieval tasks in addition to the two purchasing tasks in order to familiarize them with the interface and to conceal the purpose of the study. The tasks included searches for boot prices, prices and average lifetimes of light bulbs, and the prices and available sizes of tote bags. After two information retrieval tasks, participants used Privacy Finder to find websites offering either the sex toy or the batteries and purchased these products. The order in which participants purchased these two items was assigned randomly. The instructions specified the search strings to use to find these products. Unbeknownst to participants, these search strings returned our static search results.

Participants conducted additional information retrieval tasks between the first and second purchases. If they had purchased the batteries first, they purchased the sex toy second, and vice versa. After the second purchase, participants completed an online exit survey that asked questions about their purchases and overall reactions (Appendices J). They were required to use their own credit card and billing information for both purchases so that they would treat the purchases as "real" purchases. However, we allowed them to ship unwanted items to our laboratory. To prevent gaming of the study, we gave participants $10 in cash for completing the laboratory experiment and then another $40 by mail once we had confirmation that their orders had been shipped.[6]

6.3 Analysis

Our most significant finding was that the timing of privacy indicator display had a highly significant impact on the behavior of participants who chose to make a purchase on the first website they visited. Those participants paid for increased privacy only when their search results were annotated with privacy indicators; participants who saw the indicators at a later time were significantly more likely to ignore them. Participants who chose to comparison shop by visiting several websites before making a purchase were influenced by the privacy indicators regardless of when they were displayed. Likewise, participants' reliance on the privacy indicators also depended on whether or not they were purchasing the privacy-sensitive item, as well as the strength of the privacy indicator to which they were exposed.

In this section we describe how purchasing behaviors changed when participants were exposed to privacy indicators. Next, we examine how privacy concerns and purchasing behaviors varied based on the type of product being purchased. Finally, we detail how the timing of the privacy indicators resulted in very nuanced behaviors regarding the prices participants paid for the items, how website content had less of a role than we expected, and how timing had an impact on the number of websites participants visited.

[6] We asked participants to mail us invoices or email us tracking numbers for their purchases so that they would not plan to cancel their orders after they left our laboratory (which would make item prices less of a factor since they would not actually pay for them).

Condition	Battery Premium	Sex Toy Premium
Handicap	\$0.15	\$0.11
Privacy	\$0.34	\$0.52
Frame	\$0.26	\$0.41
Interstitial	\$0.39	\$0.49

Table 6.3: The average privacy premiums paid for both products across all four study conditions. This is the amount paid above the \$15.50 base price for increased privacy.

6.3.1 General Effects of Privacy Indicators

Hypothesis 1: Participants will pay for increased privacy when they see privacy indicators.

Attention Switch and Maintenance are the first stages of the C-HIP model where indicators may fail [140]; if users do not notice an indicator, they are unlikely to alter their behavior. In our experiment, participants reported whether they noticed the indicators during the exit survey, and we observed no significant differences between the conditions: 87.6% of participants (78 of 89) reported seeing the indicators. This number stands in contrast to studies of chrome-based security indicators, where very few participants reported noticing the indicators when they were not depicted alongside the website content [146, 135, 110]. Thus, placing contextual indicators alongside website content may increase the likelihood that the indicators are noticed and acted upon.

We compared the average price paid by participants in the control (*handicap*) condition with the average price paid by participants in the three experimental conditions to determine whether participants would pay more to shop at sites with privacy indicators than they would to shop at sites with irrelevant green indicators. We performed an ANOVA to compare the prices paid for each product between each of the experimental groups and found that when purchasing the sex toy, participants in the three experimental groups paid significantly more than participants in the *handicap* condition ($F_{3,85} = 7.938$, $p < .0005$). However, while participants in the experimental groups also paid more for batteries than those in the *handicap* condition, we did not observe any significant differences in price paid for batteries between the conditions. We concluded that participants were influenced by privacy indicators rather than by irrelevant indicators. Table 6.3 shows the average premium that participants paid for each product across all four conditions.

Our observed data corroborated the exit survey data: participants who did not see privacy indicators were less likely to consider privacy when making their purchases. We provided participants a text box on the exit survey to enter the biggest factor that they considered when making each purchase. In the *handicap* condition, 82% of participants indicated price was the primary factor during the battery purchase, and 86% indicated price for the sex toy purchase. At the same time, 9% said the website rating was the primary factor during the battery purchase, and 14% mentioned it for the sex toy purchase. In the other conditions, participants claimed price had a less important role, and the website rating was more important. In the *privacy* condition, 64% mentioned price for the batteries (36% cited the privacy rating), but only 36% mentioned price for the sex toy (55% cited the privacy rating); in the *frame* condition, 64% mentioned price for the batteries (18% cited the privacy rating), but only 46% mentioned price for the sex toy (36% cited the privacy rating); in the *interstitial* condition, 52% mentioned price for the batteries (35% cited the privacy rating), while 44% mentioned price for the sex toy (48% cited the privacy rating). As expected, when price played less of a role, the privacy ratings played more of a role in participants' purchasing decisions.

We tried to control the study by only selecting vendors that we believed would be unfamiliar to participants. During the exit survey three participants (3.4% of 89) disclosed that they had done business with our vendors in the past (two sex toy vendors and one battery vendor). However, when we asked them if previous experiences with a particular company were factors (using a 7-point Likert scale) for either purchase, we found no correlation between self-reported familiarity and where participants made purchases during the study.

In the exit survey, we asked participants to define what they believed the indicators represented in order for us to determine whether those in the experimental conditions understood that they corresponded to privacy levels, and whether those in the *handicap* condition understood they represented handicap accessibility. We discovered that 68.2% of those in the *handicap* condition (15 of 22) correctly understood the meaning, 81.8% of those in the *privacy* condition (18 of 22), 63.6% of those in the *frame* condition (14 of 22), and 52.2% of those in the *interstitial* condition (12 of 23). We can see that comprehension rates were higher among those who saw the indicators as search result annotations, however these differences were not statistically significant.

After performing a t-test on the the prices paid versus whether participants noticed the indicators, we discovered that when the indicators represented privacy—that is, in all conditions except for the *handicap* condition—participants who noticed the indicators paid significantly more money for both the batteries ($t_{65} = 3.026$, $p < 0.004$) and the sex toy ($t_{65} = 2.569$, $p < 0.012$). This was not significant when the indicators represented handicap accessibility; participants who noticed the handicap accessibility indicators were unlikely to spend more money than those who did not notice them. Unfortunately, we were unable to yield statistically significant results when comparing indicator understanding—whether participants could correctly identify the indicators as representing privacy—with the amount that participants spent. This may be because over two thirds of the participants were able to correctly identify the indicators, regardless of whether or not they noticed them during the study, and because understanding was self-reported whereas purchase decisions were observed. However, clearly labeling the meaning of the indicators and displaying them at the user's locus of attention may minimize errors during the *Comprehension and Memory* stage of the C-HIP model [140].

6.3.2 Product-Specific Privacy

Hypothesis 2: Participants who see privacy indicators will pay more for the privacy-sensitive item than the item that does not raise additional privacy concerns.

We performed a pairwise t-test across both purchases to compare the prices paid for the sex toy with the prices paid for the batteries in each condition (Table 6.3), and found that participants paid significantly more—for higher privacy levels—for the sex toy than for the batteries in both the *privacy* ($t_{21} = 2.935$, $p < 0.008$) and *frame* ($t_{21} = 2.346$, $p < 0.029$) conditions.

What we found most interesting was that participants in the *interstitial* condition did not pay significantly more for one product versus the other. Instead, they paid a privacy premium for both products. In this case, the effect of the privacy indicators being displayed as an interstitial (i.e. a more active type of indicators that likely forced users to revisit their decisions of which website to visit from the search results) diluted the role of product-specific concerns when the participants made their purchases. Thus, they were motivated to find the high privacy websites for both products.

We compared our observed data to the self-reported data that participants provided on our exit survey.

Information	$\mu_{sex\ toy}$	$\mu_{battery}$	t_{88}	p-value
Credit card	4.92	4.55	2.938	.004
Email address	4.87	3.96	5.002	.0005
Physical address	4.29	3.45	4.738	.0005
Phone number	4.62	3.94	4.008	.0005
Purchase history	3.87	2.92	5.499	.0005

Table 6.4: Participants used a 7-point Likert scale to specify how concerned they were during each purchase when providing various types of personal information.

In the exit survey we asked participants to rate their privacy concerns for both products on a 7-point Likert scale (six represented "extremely concerned," while zero represented "not concerned at all"). Participants reported an average concern level of 5.56 for the sex toy ($\sigma = 2.291$) and 3.56 for the batteries ($\sigma = 1.864$). We performed a paired t-test and determined that participants had significantly higher levels of concern when purchasing the sex toy ($t_{88} = 7.884$, $p < .0005$). Participants used another 7-point Likert scale to specify how concerned they were during each purchase when providing specific types of information: credit card numbers, email addresses, physical addresses, phone numbers, and purchase histories. For each piece of information, participants were significantly more concerned about what would happen to that information when they provided it for the sex toy purchase than for the batteries purchase, as shown in Table 6.4.

Participants who saw privacy indicators were able to address many of their privacy concerns by purchasing the sex toy from websites with better privacy policies. However, this was not the case for those in the *handicap* condition, who did not see the privacy indicators.

We were initially concerned that some participants may have behaved in a predictable fashion because of the Hawthorne Effect; they may have made purchases from high-privacy websites because they understood that that was what we were examining. To examine this effect, we examined if the order in which the products were purchased had an effect on where participants made their purchases. That is, if participants purchased the sex toy first, it may have raised their privacy concerns to the point that they were more observant of the privacy indicators during the subsequent battery purchase (i.e. priming). However, since we did not observe any statistically significant differences in behavior based on the order in which products were purchased, we concluded that this effect was likely not present among our sample population. Similarly, we were concerned that those who shipped items to our laboratory may have done so to mitigate the effects—in their minds—of making purchases from cheaper/low-privacy merchants. Seventeen participants sent batteries to our laboratory, while thirty-four sent us sex toys. We were surprised to discover that not all of those who shipped us batteries shipped us sex toys. Because we also did not observe statistically significant correlations between purchase decisions and whether they sent us their items, we concluded that privacy was not a motivation for sending us the items. Instead, it is likely that participants chose to send us items because they either did not want them or because they believed they would be reimbursed faster that way.

6.3.3 The Effect of Timing on Prices

Hypothesis 3: Participants will be more likely to pay for increased privacy when they see privacy indicators alongside search results before visiting a website than when they see privacy indicators after clicking on search result links.

Condition	Websites	Batteries	(n)	Sex toy	(n)
Handicap	1	\$0.16	(13)	\$0.10	(16)
	>1	\$0.14	(9)	\$0.17	(6)
Privacy	1	\$0.41	(14)	\$0.46	(13)
	>1	\$0.22	(8)	\$0.61	(9)
Frame	1	\$0.03	(8)	\$0.06	(8)
	>1	\$0.39	(14)	\$0.61	(14)
Interstitial	1	\$0.03	(8)	\$0.19	(8)
	>1	\$0.58	(15)	\$0.65	(15)

Table 6.5: Average privacy premiums paid—above the base price of \$15.50—for each product by participants in the four study conditions. The study conditions are broken down based on whether participants visited multiple websites before making a purchase. The numbers in parentheses reflect the size of the groups.

Hypothesis 4: Participants will be more likely to pay for increased privacy when they see privacy indicators before they see the content of a website than when they see privacy indicators alongside the content of a website.

The results of our study indicate that the impact of timing was nuanced: Hypothesis 3 was correct for participants who clicked only one search result, but false for participants who visited multiple websites before deciding where to purchase. Table 6.5 shows the average prices paid for each product across the four study conditions, broken down based on whether participants visited more than one website.

One-click purchases

We performed an ANOVA to compare the amounts participants paid between the different conditions when they visited only one website before purchasing the batteries ($F_{3,39} = 4.772$, $p < 0.006$). We discovered that participants in the *privacy* condition paid significantly more than those in the *frame* ($p < 0.019$) or *interstitial* ($p < 0.019$) conditions.[7] This indicates that participants used the search result annotations to choose websites with increased privacy levels. However, when the privacy indicators were displayed after participants had selected websites from the search results, the participants ignored those indicators, perhaps because they were unwilling to return to the search results. Instead, they were focused on the purchasing task. For these participants the increase in privacy for the batteries was not worth the hassle of selecting new websites from the search results.

We observed slightly different behaviors when participants purchased the sex toys. Again, we observed significant differences between the study conditions ($F_{3,31} = 4.402$, $p < 0.009$), but now the differences were between the *privacy* condition and the *handicap* ($p < 0.012$) and *frame* ($p < 0.027$) conditions. Again, participants in the *privacy* group paid more for privacy when visiting only one website because they saw the privacy indicators before choosing a website to visit. The lack of a significant difference between the *privacy* and *interstitial* conditions is likely a random phenomenon that may disappear with a larger sample size.

Recall that all our study participants claimed to have a desire for privacy. However, several of the participants were still unmotivated to purchase from the high-privacy websites when the privacy indicators

[7] All post-hoc analysis throughout this paper was done using Tukey's HSD test.

were displayed after they had already made a decision to visit particular website. At the same time, this error in the *Motivation* stage of the C-HIP model was mitigated when the privacy indicators were displayed alongside the search results, before the participants had made up their minds about which website to visit.

Multiple-click purchases

Of the participants who visited multiple websites before purchasing an item, we found that the timing of the privacy indicators did not significantly impact the selection of the website from which they made their purchases. An ANOVA yielded significantly different prices paid for the batteries between the study conditions ($F_{3,42} = 5.424$, $p < 0.003$). Using post-hoc analysis we discovered that participants in the *interstitial* condition paid significantly more than participants in both the *handicap* ($p < 0.004$) and *privacy* ($p < 0.030$) conditions. However, there were no significant differences in battery prices when comparing the *frame* condition with the *handicap* and *privacy* conditions. This can likely be attributed to the role of website content—those who viewed content alongside the privacy indicator relied on the privacy indicator less. It is also likely that because the interstitial interrupted their immediate task and required their attention to dismiss it, the strength of this privacy indicator was greater than that of the other two.

The significantly stronger effect of the *interstitial* condition was only observed during the battery purchase: we observed significant differences between the conditions when examining prices paid by participants who visited multiple websites when purchasing the sex toy ($F_{3,40} = 8.860$, $p < 0.0005$), but this was because everyone exposed to privacy indicators—regardless of timing and placement—paid significantly more than those in the *handicap* condition ($p < 0.001$ for *handicap* vs. *privacy*, and $p < 0.0005$ for both *frame* and *interstitial* vs. *handicap*). This is interesting because it means that those who saw privacy indicators after choosing websites from the search results still ended up purchasing the sex toy from the higher privacy websites—it just took them longer to find them.

6.3.4 The Effect of Timing on Website Visits

Hypothesis 5: Participants who see privacy indicators after clicking on search result links will visit more websites than those who see privacy indicators alongside search results.

We further explored the role of timing by examining the number of search results visited by participants in the *frame* and *interstitial* conditions. Recall that these participants only saw privacy indicators after selecting search results. Table 6.6 shows the number of websites participants in all conditions visited on average before making a purchase. We performed an ANOVA and found significant differences between the conditions for both the battery ($F_{3,85} = 4.475$, $p < 0.006$) and the sex toy ($F_{3,85} = 8.394$, $p < 0.0005$) purchases.

Because we were primarily interested in how long it took participants to find the websites with the highest privacy levels, we performed another ANOVA, though this time only examining participants who purchased from the websites with four green boxes. When purchasing the batteries, participants in the *privacy* condition clicked significantly fewer search results to find the website with the four green boxes ($F_{3,22} = 23.126$, $p < 0.0005$). Participants in the *interstitial* and *frame* conditions clicked 203% more search results on average than those in the *privacy* condition to purchase from this same website and obtain the same level of privacy ($p < 0.0005$ for both comparisons). Thus it took participants in the *interstitial* and *frame* conditions significantly longer to find the same high-privacy website that those in the *privacy* condition were able to locate with a single click.

Condition		Batteries	Sex Toy
Handicap	(22)	1.86 ($\sigma = 1.17$)	1.41 ($\sigma = 0.91$)
Privacy	(22)	1.86 ($\sigma = 1.36$)	1.73 ($\sigma = 1.12$)
Frame	(22)	3.05 ($\sigma = 1.79$)	3.09 ($\sigma = 1.77$)
Interstitial	(23)	3.09 ($\sigma = 1.78$)	3.04 ($\sigma = 1.69$)
Interstitial*	(23)	2.09 ($\sigma = 1.38$)	1.74 ($\sigma = 1.10$)

Table 6.6: The total number of search results visited (out of a maximum of five) before participants purchased each product. The last row shows the number of sites visited by members of the *interstitial* condition when they chose to proceed to the website in light of the privacy indicator.

Recall that in the *interstitial* condition, participants must acknowledge the privacy indicator before viewing the destination website. If instead of examining the number of search results clicked, we examine the number of websites viewed by those in the *interstitial* condition, we no longer see a significant difference between the *interstitial* condition and the *privacy* and *handicap* conditions. That is, when participants encountered the interstitial privacy indicator on a website with a low privacy level, they were more likely to return to the search results without viewing that website.

This distinction was also apparent when we examined the number of search results clicked prior to purchasing the sex toy from the website with the highest privacy level ($F_{3,33} = 21.039$, $p < 0.0005$): participants in the *interstitial* and *frame* conditions clicked an average of 168% more websites ($p < 0.0005$ for both comparisons) than those in the *privacy* condition. Again, participants in these three conditions did not differ on the level of privacy they achieved, it merely took them longer to achieve that same level of privacy when the indicators were displayed after search results were selected. Therefore, displaying privacy indicators alongside search results creates more efficient shopping experiences for most users, while also helping users who click fewer search results to achieve greater levels of privacy.

6.3.5 Limitations & Future Work

While we demonstrated that the timing of a privacy indicator's appearance has an impact on whether users visit websites with better privacy policies, there are still many unanswered questions. We did not compare the effect of privacy indicators with other relevant indicators such as customer ratings, nor did we explore the extent to which participants might view privacy indicators as a proxy for other indicators of trustworthiness unrelated to privacy. Two additional areas that we plan to focus on in future studies are how consumers make decisions about privacy premiums and how website content competes with indicators for a user's attention.

Privacy Premiums

We observed that participants were willing to pay premiums to receive higher levels of privacy. In this particular study we used a privacy premium of $0.75. However, we do not know if participants view privacy premiums as a percentage of a purchase price or as a flat rate. That is, would participants have paid this same premium on an item that cost half as much? Would participants pay a $1.50 privacy premium on an item that cost twice as much?

Website Content

Fogg et al.'s work on website credibility indicates that the "look and feel" of a website is the main factor when users make trust decisions [55]. However, we were surprised to discover that this was not always the case: many times participants placed more weight on the privacy indicators than the websites. That being said, it is unclear how exactly participants assessed the quality of the websites they visited. Future studies might examine how participants assess the look and feel of websites while also examining their reactions to privacy indicators.

Habituation

Additionally, we are unsure of the long-term effects of interacting with our privacy indicators. That is, will Privacy Finder users eventually become habituated and stop paying attention to the indicators? We test this in the next chapter by performing a field study of Privacy Finder users over ten months.

6.4 Conclusion

In this study we showed that the timing and placement of how privacy indicators are displayed impacts purchasing decisions: participants who decided to visit only one website to make their purchases paid significantly more money for a higher level of privacy when privacy indicators were presented alongside their search results; similar participants who did not see privacy indicators until after they had already selected a website were unwilling to spend time finding websites with higher privacy levels and instead made purchases from cheaper websites. Likewise, participants who did comparison shopping were just as willing to use interstitial and frame privacy indicators to find websites with higher privacy levels, even though this meant visiting significantly more search results.

By examining this study through the framework of the C-HIP model, we were able to create recommendations for improving contextual indicators that address common errors in the *Attention Switch and Maintenance*, *Memory and Comprehension*, and *Motivation* stages.

Finally, we observed that privacy decisions depended on privacy concerns surrounding the items being purchased: participants had greater privacy concerns when making the sex toy purchases and therefore went out of their way to use the privacy indicators to find websites that offered higher levels of privacy, even if this meant paying a premium. Likewise, many participants were not willing to pay a privacy premium for the batteries because the product did not trigger the same level of privacy concern as the sex toy.

Chapter 7

Privacy Finder Usage Study

This chapter is largely a reproduction of a paper co-authored with Janice Tsai, Lorrie Faith Cranor, and Alessandro Acquisti [128]. This research was funded in part by U.S. Army Research Office contract no. DAAD19-02-1-0389 (Perpetually Available and Secure Information Systems) to Carnegie Mellon University's CyLab and by Microsoft Research.

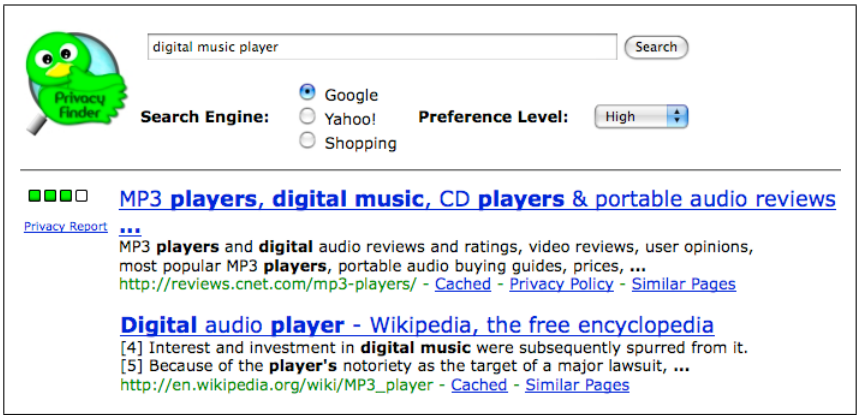


Figure 7.1: The Privacy Finder search interface.

In the previous chapter I presented various designs for web browser privacy indicators. I presented results from a laboratory study that indicated that search result annotations benefited the most users (Figure 7.1) [49]. I also showed how contextual indicators could be designed to minimize errors in the *Attention Switch and Maintenance*, *Comprehension and Memory*, and *Motivation* stages of the C-HIP model. To validate these results, we conducted a field study over the course of ten months, recruiting 460 participants to use Privacy Finder for their online searches. We found that privacy indicators that appeared in search results impacted browsing patterns. Results with privacy indicators experienced higher visitation rates as compared to sites without indicators, even when the sites with privacy indicators were positioned lower on the search results page. In fact, participants were significantly more likely to visit search results further down on the results pages when those results were annotated with privacy indicators, and sites with the highest privacy ratings had the highest overall visitation rates. On the other hand, sites with low privacy ratings did

not have significantly different visitation rates than sites without any privacy ratings. Finally, we examined these indicators with regard to the *Attitudes and Beliefs* and *Behavior* stages of the C-HIP model.

7.1 Methodology

One of the motivations behind the development of Privacy Finder was the notion of making privacy information more prominent. When privacy information is made more accessible or evaluable, it may be more likely that people will consider the level of privacy protections afforded by a specific site when selecting a website from lists of search results. We designed our field study of Privacy Finder to test whether users do, in fact, take privacy information into account. Specifically, we tested the following hypotheses:

- By displaying privacy information alongside search results, users will be more likely to visit websites that offer higher levels of privacy protection, as denoted by our privacy indicators.
- By displaying privacy information in the search engine, users will be more likely to visit websites further down the list of search results when those sites have privacy indicators, as compared to visitation rates when no privacy indicators are present. Sites with privacy indicators will have a higher probability of being visited than sites in the same position without privacy indicators.

7.1.1 Recruitment

From December 2007 to October 2008, we recruited participants to test a privacy-enhanced search engine. We posted announcements about the study on various volunteer solicitation websites, including Craigslist and online sweepstakes sites. We used a raffle as an incentive for people to use our search engine regularly. For each day that participants conducted searches using Privacy Finder, they were issued a raffle ticket. We conducted weekly raffles of $20 Amazon.com gift certificates, and a $200 Amazon.com grand prize raffle.

To participate in the study, participants first registered their email addresses and completed a pre-study survey that contained questions about their attitudes toward online privacy (Appendix K). Subsequently, when participants logged in to Privacy Finder using their email addresses, we placed a cookie on their computers. The cookie was used to distinguish the study participants from other Privacy Finder users and to create entries in our prize drawing database. All searches were anonymized, and we used the email addresses solely to contact participants in the event that they won a prize drawing. When users were logged in (i.e. we detected the cookie), we recorded their queries, the times and dates of the queries, the search engine selected, the privacy level of the search engine, the search results that were returned, the privacy ratings for those search results, and any results visited.

7.2 Data Analysis

We analyzed our pre-study survey results to gain information about the population of users who participated in our study and to understand their levels of privacy concern. Then, we analyzed our Privacy Finder search data by comparing the browsing behavior of users whose queries produced search results that contained privacy indicators—websites with P3P—with the behavior of users whose queries produced a set of search results without privacy indicators. By using statistical tests to compare visitation-rates between these different types of queries, we investigated whether privacy indicators impacted browsing behavior.

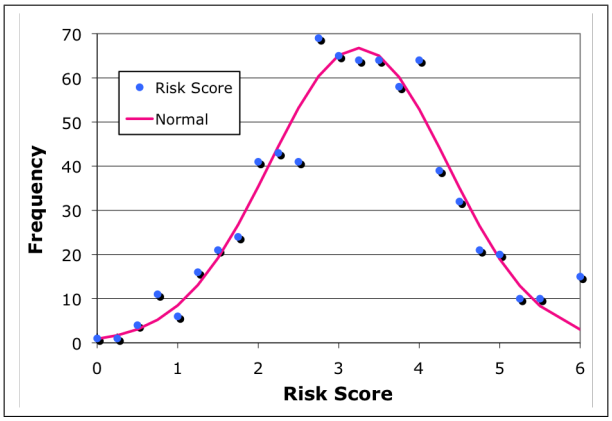


Figure 7.2: The histogram for the risk scores for our participants as compared to the normal distribution, plotting the risk score and the number of people who had that same risk score. We see that the risk scores have a good fit to the normal distribution, bin size 0.25.

7.2.1 Pre-study Survey

During the participant recruitment process, we asked potential subjects to complete a 10-question online survey. We collected survey responses from 740 people.[1] The average age of our participants was 34.7 years, and 57.7% of the respondents were female. Our sample was also relatively highly educated, with 88.2% of respondents having a college education. Based on a 7-point Likert scale from "Never" (1) to "Always" (7), we found that, on average, participants "sometimes" noticed whether or not a website has a privacy policy (μ = 3.89, 95% CI = 3.76 - 4.03), and that they do not often read website privacy policies (μ = 2.82, 95% CI = 2.70 - 2.94).

In addition to collecting basic demographic information, we queried our respondents about their privacy concerns. We used a four-item risk belief scale developed in previous studies to calculate a risk score for each participant [129]. Participants' responses to the four 7-point Likert scale questions were averaged (the lower the score a respondent receives, the less concerned they are about their privacy). The risk belief scale consists of the the following questions:

- I feel safe giving my personal information to online stores. (Strongly disagree to strongly agree (*reversed*))
- Providing online stores with personal information involves too many unexpected problems. (Strongly disagree to strongly agree)
- I generally trust online companies with handling my personal information and my purchase history. (Strongly disagree to strongly agree (*reversed*))
- How concerned are you about threats to your personal privacy online in American today? (Not concerned at all to extremely concerned)

[1] Only 62% of those who completed the survey chose to participate in the search result analysis portion of the study.

We plotted a histogram of participants' risk scores, as shown in Figure 7.2, where the bin size is 0.25. This histogram and the superimposed normal distribution curve (μ = 3.25, σ = 1.11) indicate that the risk scores for our sample are well represented by a normal distribution.[2] This distribution indicates that a majority of our participants had a medium level of privacy concern, with a slightly higher proportion of higher concern respondents than low concern respondents.

The results of our pre-study survey mimics those found in the Westin surveys. Alan Westin conducted a series of privacy concern surveys that gathered longitudinal data about the level of privacy concern and online privacy concern among Americans. In 1996, he created a "Privacy Concern Index" that divided respondents into three categories: the privacy fundamentalist (high privacy concern), the privacy pragmatist (medium privacy concern), and the privacy unconcerned (low privacy concern) [134]. In this 1996 survey and in subsequent Westin surveys, we see that the majority of respondents are classified as privacy pragmatists, with a slightly higher population of privacy fundamentalists than the privacy unconcerned [80].

Our sample of subjects may suffer, naturally, from self-selection bias, albeit of a particular nature: our subjects were drawn to participating in a study which explicitly focused on privacy protection; yet they were also willing to reveal (albeit anonymously) their search results to the researchers.[3] However, we created a monetary incentive to recruit participants, which also served to counter-weight the potential biasing effect of the privacy incentive. Specifically, we offered a weekly raffle incentive to keep people interested in the search engine and to promote its use. Based on the responses to questions in the pre-study survey (see Figure 7.2) we conclude that, in fact, we did not only attract individuals who were highly concerned about their privacy: rather, we see that the majority of our respondents had medium levels of privacy concern. Thus, even though our sample was self-selected, their privacy concerns are likely representative of the larger population.

7.2.2 Experimental Control

When designing a research study, researchers must always consider how they will design or deploy an experimental control. In this field study, we were asking users to use Privacy Finder as their normal search engine. We would have ideally liked to create a control for Privacy Finder (a search engine that did not annotate search results with privacy indicators) or a method to test other indicators (e.g. a search engine that annotated results with merchant rating indicators). By assigning participants randomly to these additional conditions, we would have been able to directly compare the behavior of users under identical conditions but without privacy indicators, with privacy indicators, and with other indicators. A large search engine operator could easily setup such conditions by simultaneously deploying privacy indicators and another type of indicator each to a small subset of their users, while continuing to provide no indicators to most users. However, this is much more difficult to setup when no existing users are available and new users must be recruited for each experimental condition.

After considering the use of these additional search engine conditions, we determined that it would be too difficult to find enough users to use several different search engines for long periods of time without offering a good reason why. In the commercial search engine market, a small number of search engines have

[2] Using Pearson's chi-squared test, we see that we cannot reject the null hypothesis that the risk scores for our participants are consistent with the normal distribution, $\chi^2 = 0.36$.

[3] Subjects were told that their searches would be logged, but that we would not individually identify them, other than to inform them in the event they won a prize.

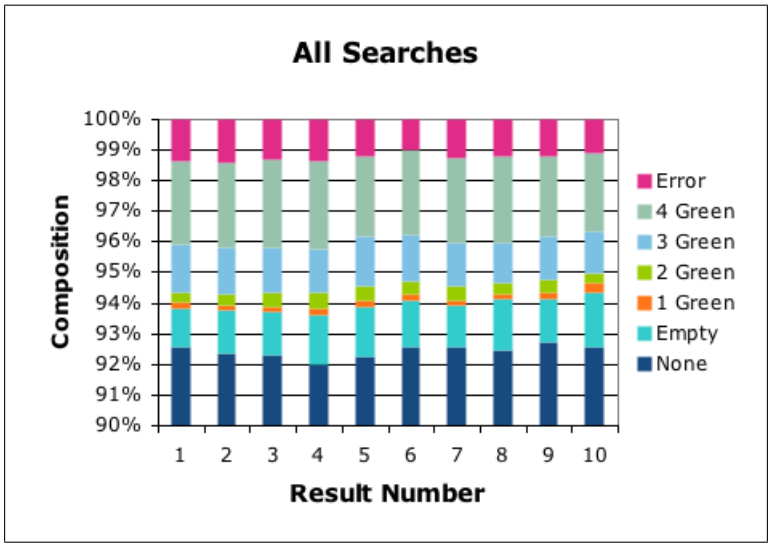


Figure 7.3: Composition of search results based on privacy ratings and position on the search results page.

Privacy Indicator	# of Results	% of results with indicator
No Indicator	134,340	92.43%
0 Green	2,181	1.50%
1 Green	289	0.20%
2 Green	595	0.41%
3 Green	2,125	1.46%
4 Green	4,003	2.75%
Error	1,807	1.24%

Table 7.1: The frequency with which each privacy indicator appeared in the search results.

maintained the majority of search engine market share for over half a decade. The majority of searches are conducted on the Google (63.0%), Yahoo! (21.0%), or Microsoft (8.3%) sites, indicating that it is difficult to grab market share from the large search engine players [27]. In the meantime, many search engines have gone out of business [121]. Despite the low switching costs of choosing a new search engine [60], it seems that the most successful search engines provide such good quality results or have added features (e.g. integration with email) that it makes users reluctant to switch to new search engines.

Instead of implementing a control search engine condition, we used a within study control for statistical analysis purposes. We partitioned the search results pages returned to users into two sets: those with at least one privacy indicator and those with no privacy indicators (because none of the results on those pages had P3P). We compared the visitation rates of results in the privacy-indicator set with those in the no-privacy-indicator set as the control.

# of Results with Indicator	# of searches	% of searches
No Indicators	9481	62.72%
1	3,102	20.52%
2	1,348	8.92%
3	544	3.60%
4	244	1.61%
5	175	1.16%
6	73	0.48%
7	42	0.28%
8	31	0.21%
9	21	0.14%
10	55	0.36%

Table 7.2: The frequency of results pages annotated with 0-10 privacy indicators. For example, there were 55 pages where all 10 search results were annotated with privacy indicators.

7.2.3 Privacy Finder Usage Data

Over the course of the study, 460 unique users logged in and allowed us to track their searches. These users conducted 15,116 queries over a ten month period. On average, each participant used Privacy Finder for six days and conducted 33 queries, with a median of four queries.

Privacy Finder allowed users to select which search engine they wished to use (Google, Yahoo!, or Yahoo! Shopping) and to customize the level of the privacy preference setting (low, medium, high, or custom). Google was the default selection for the search engine, and was used for over 80.70% of the searches. Yahoo! was used for 18.34% of the searches (2,633) and Yahoo! Shopping was used for 0.96% of the searches (147). The majority of searches were made at the default privacy setting of *medium* (91%), with 5% of the searches made using the *high* setting, 3% at *custom*, and 1% at *low*.

Privacy Finder computes a privacy rating based on elements of websites' privacy policies and the privacy preferences setting in Privacy Finder. The frequencies with which each privacy indicator appeared in the search results are depicted in Table 7.1. Most searches returned a page with ten search results, although some queries returned fewer results. The queries conducted in this study returned a combined total of 145,340 search results, of which 6.33% were annotated with privacy indicators and 1.24% were P3P-enabled but no privacy rating could be computed due to errors in their P3P policies.

We examined the frequency of search results with privacy indicators. We found that the majority of queries returned results without any P3P policies, and therefore without any privacy indicators. The frequencies of the number of privacy indicators per set of search results are summarized in Table 7.2. The highest privacy rating (four green boxes) was the privacy indicator that occurred most frequently in our data set. As shown in Figure 7.3, the frequencies of each of the privacy indicators were evenly distributed across all pages of ten search results. (A chi-square test shows that there are no statistically significant differences in the distribution of P3P-rated results by result number, $p = 0.47$, $\chi^2 = 8.64$.) This indicates that it was not the case that a specific result number was more likely to be annotated with a privacy indicator (i.e., participants are not more likely to visit result 3 simply because there were a higher number of search results with

privacy indicators that happened to be in position three). On average, each of the ten search result positions was annotated with a privacy indicator 7.57% of the time.

We categorized the search terms participants used to determine the types of queries conducted (navigational, transactional, or informational). We found that the most frequent searches were navigational in nature. Nine out of the top ten searches were navigational. The top ten searches made up about 1% of the total queries conducted. When users conduct navigational queries, they typically know which website they are looking for. We found that when participants visited search results for the nine navigational queries in our top ten most frequent searches, they visited the first result 79.7% of the time.

In our data analysis, we examined the position of each search result and how frequently search results were visited. A "visit" in the context of this paper refers to the user clicking on a website in a set of search results in order to go to that specific website. Examining our dataset for usage changes between the search engines, we found no statistical differences in browsing patterns between the Google and Yahoo! search engines (when using a chi-square test to examine the proportion of privacy-annotated results visited). Due to the small sample size of the Yahoo! Shopping searches, we eliminated those searches from the remainder of the analysis. We also filtered out searches where none of the search results were visited. Our hypotheses focus on browsing behavior; searches without clicks are irrelevant to answering our research questions.

Our final dataset consisted of 7,046 queries made through the Google and Yahoo! search engines where at least one search result on the search results page was visited. Of these queries, 79.1% were made through Google (5,571) and 20.9% (1,475) through Yahoo!.

7.2.4 Browsing Patterns

Throughout the study, we found it difficult to retain users. We found that people were mostly likely to sign up for the study, conduct multiple searches over the course of the day, and fail to return to the Privacy Finder site on subsequent days (294 out of 460 participants). To determine if the browsing patterns of these users (1,030 queries) skewed our results (due to the novelty of seeing privacy indicators), we compared the proportions of visits to P3P-rated sites for the *one day* users to the visitation patterns of users who participated in the study over a longer period of time. We found that *one day* users visited 5.22% (46 out of 881) of the search results that had privacy indicators while the rest of the participants in the study visited 8.27% (825 of 9,975) of the search results with privacy indicators, Fisher's exact $p < 0.001$. This indicates that the privacy indicator-annotated search result visitations were significantly different, but that *one day* users were actually *less* likely to visit sites with privacy indicators. Despite the novelty of privacy indicator annotated search results, these indicators did not significantly sway their search result visitations.

To further determine if continued use of Privacy Finder would alter search behavior over time (i.e. would users become habituated to the privacy indicators), we examined the search queries of the 32 participants who used Privacy Finder for two weeks or longer. We specifically examined the searches made over the first seven days of participation and compared them to the searches made over the second set of seven days. We find that for the first seven days of study participation, these participants visited 7.74% (121 of 1,563) of the search results that were annotated with privacy indicators. Over the second week, these participants visited 7.96% (93 of 1,168) of the sites with privacy indicators. These proportions of visitations were not statistically different (Fisher's Exact $p = 0.83$). It appears that continued use over this 14-day period did not significantly alter browsing behavior; participants continued to visit sites with privacy indicators at about the same rate. This indicates that within this sample, participants did not become habituated to the indicators,

which is likely because those who continued to use Privacy Finder for an extended period of time did so because they saw value in the indicators (i.e. they found the privacy ratings helpful). Thus, by providing contextual indicators that users both notice and care about, we were able to address the issue of habituation, which is a common pitfall of the *Attitudes and Beliefs* stage of the C-HIP model [140].

7.2.5 Data Validation

In addition to our analysis of the privacy attitudes of the participants in the study and the within-study control, we also validated the search result visitation rate in our dataset. We were interested in knowing whether or not people were visiting search results at a normal rate or if they were attempting to falsify their visitation patterns (i.e. visiting the last search result in all of their search queries).

To address potential concerns with our data, we validated the use of Privacy Finder to search behavior data from a major search engine. Microsoft provided us with the Spring 2006 search data collected from users of their search engine, Live Search.[4] This data consisted of about 15 million queries sampled over one month. The attributes of this dataset included query strings, timestamps, any URLs visited, and the positions on the results page for each URL visited. This dataset only contained queries where a user visited at least one of the URLs. We will refer to this dataset throughout the paper as the *MS Live* dataset.

We can determine the likelihood that users will visit a certain website based on its position on the results page. Based on the methodology implemented in the study of search results by Agichtein et al. [8], we calculated the relative click frequencies for the *MS Live* and *Privacy Finder* search results. This allows us to evaluate the proportions of visitations relative to the first search result. This provided us with a standardized method with which to evaluate search result visitations across the two datasets.

We calculated the relative visitation frequencies in two parts. First, the actual frequency of visiting a search result was calculated for each result position. Second, these frequencies were normalized by comparing them to the first result so that the relative frequency of a visit at the top position was 1.0. We calculated the relative click frequency for the two datasets, and compared the *MS Live* click frequency to the click frequency for the *Privacy Finder* dataset. Figure 7.4 shows that the relative click frequencies for the *Privacy Finder* and *MS Live* datasets were very similar.

The search patterns from our study participants seem to mimic those from a real world search engine. There was a 50% overlap in the top ten search terms in these two data sets. The top five search terms common to the two data sets were “Google,” “Yahoo,” “Amazon.com,” “eBay” and “MySpace.” This is reassuring, in that it appears that our users were not focused on the privacy indicators, but on conducting real search queries.

7.3 Results

After examining these general browsing patterns, we proceeded to test our first hypothesis:

Hypothesis 1: By displaying privacy information alongside search results, users will be more likely to visit websites that have high levels of privacy, as denoted by our privacy indicators.

[4] http://www.live.com

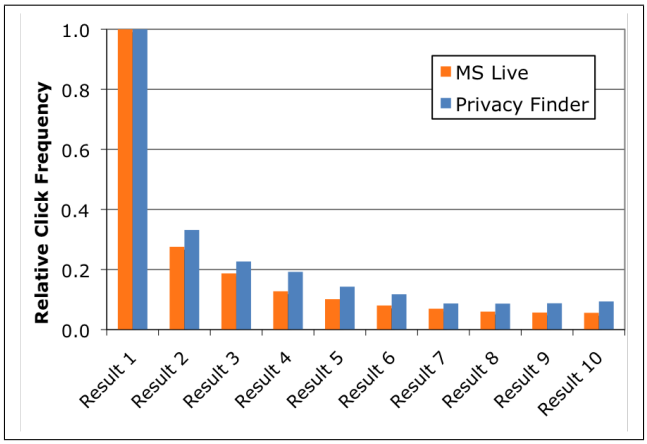


Figure 7.4: Relative click frequency rates for the *Privacy Finder* and *MS Live* datasets based on position on the search results page.

Privacy Indicator	% Results Visited	Fisher's Exact p
0 Green	13.66% (144)	0.63
1 Green	8.60% (8)	0.14
2 Green	13.21% (42)	0.68
3 Green	14.73% (137)	0.67
4 Green	17.39% (367)	$<$ **0.001**
Error	15.46% (143)	0.30

Table 7.3: Comparison of visitation rates between search results without privacy indicators (14.24%) to visits to search results annotated with privacy indicators. Significantly more users visited search results annotated with the highest privacy rating (Fisher's exact $p < 0.001$).

To test *Hypothesis 1*, we calculated the visitation rates to search results annotated with each type of privacy indicator. We calculated the probability that a user will visit a result based on its privacy rating: the chance of someone visiting a site if it has no privacy rating, 0-4 green boxes, or the P3P error icon. On average, regardless of the position on the search results page, a site without a privacy rating was visited 14.24% of the time. When a site had a high privacy rating—four green boxes—it was visited 17.39% of the time.

We compared the proportion of visits to websites with each level of privacy rating to the proportion of visits to those sites that were not annotated with any privacy indicators (14.24%). Table 7.3 shows the results of the visitation comparisons between each privacy indicator as well as the statistical significance for those proportions based on Fisher's exact test. To account for multiple tests, we applied the Bonferroni correction by setting $\alpha = 0.008$.

We found that having low or medium privacy ratings (0-3 green boxes) had no detrimental effect on visitation rates: our statistical tests indicated that there was no observable difference between the visitation rates for results annotated with low or medium privacy ratings and the visitation rates to the sites without privacy ratings. Instead, we found that having a high privacy rating—four green boxes—significantly increased the

number of visits to those sites.

To evaluate the overall impact on visitation rates to sites with privacy indicators, in general, we grouped all the search results annotated with privacy indicators. We found that sites with privacy indicators attracted a greater proportion of visits (15.49%, or 841 of 5,429 including websites with P3P errors) compared to sites without P3P (14.24%, or 9,145 of 64,221). We performed a one-tailed Fisher's exact test and found that this result was statistically significant ($p < 0.007$). We conclude that that privacy indicators have an impact on which website a user decides to visit.

To determine whether having a higher privacy rating induces people to visit search results lower on the page despite the presence of other less highly rated search results with privacy indicators, we examined the visitation patterns of sets of search results with multiple privacy indicators in a single search query. To determine if these lower-postioned high privacy websites had any impact on browsing patterns, we compared the visitation rates to sets of results with one privacy indicator to the visitation rates of sites with multiple indicators. We found that a total of 5,302 searches had exactly one privacy indicator in their set of search results. Of those searches, users visited sites with privacy indicators 416 times. This indicates that when people are presented with search results with exactly one privacy indicator, they visited that result 7.85% of the time. Comparatively, there were 311 searches that had multiple privacy indicators on a page where a site with a higher privacy rater was in a lower position on the search engine page (e.g. search result 3 had an indicator with two green boxes but search result 6 had an indicator with four green boxes). Of these cases, users visited the lower (higher-rated) result for 35 of those searches, for a proportion of 11.25%. We conclude that our participants were influenced by better privacy indicators, visiting sites with better ratings a higher proportion of the time when they were available (Fisher's Exact $p = 0.04$).

We find that *Hypothesis 1* is supported: users presented with privacy information were more likely to visit websites that had a high privacy rating. This corroborates the results of the previous chapter in this thesis, which found that people who care about privacy would visit high-privacy websites if those websites are annotated with privacy indicators. This also addresses the *Behavior* stage of the C-HIP model [140]: by providing privacy-conscientious users with search result annotations that indicate website privacy levels, the users will be more likely to visit the high-privacy websites.

Hypothesis 2: By displaying privacy information in the search engine, users will be more likely to visit websites further down the list of search results when those sites have privacy indicators, as compared to visitation rates when no privacy indicators are present. Sites with privacy indicators will have a higher probability of being visited than sites in the same position without privacy indicators.

To test *Hypothesis 2* and to examine the impact of position on the search results page and website visitations, we compared two subsets of our *Privacy Finder* dataset:

No Indicator: The *No Indicator* data acts as the control and includes all the queries whose sets of search results did not contain any privacy indicators. When users see search results that fall under this category, they would see the results without any additional indicators or privacy-related information.

One Indicator: The *One Indicator* data consists of the searches where there was exactly one search result with a privacy indicator on the search results page. This dataset controls for the effect of having a privacy indicator at a specific position on the search results page. Otherwise, the presence of multiple

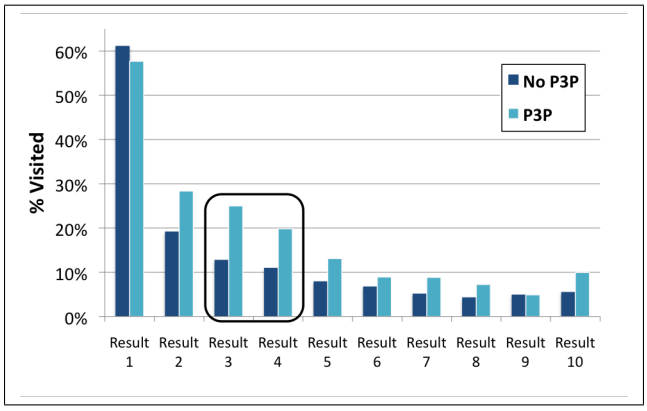


Figure 7.5: Visitation rates for the *No Indicator* and *One Indicator* search results based on the position on the search results page. The circle around Results 3 and 4 indicate that these specific search results were visited at a significantly higher rate when websites in those positions had privacy indicators.

privacy indicators on a single page may be a confounding factor, making it more difficult to examine the impact of the indicators on the probability that a user will visit a certain result based on its position.[5]

We compared the proportions of visitations for each result position for the two datasets. For each result position, a Fisher's exact test was used to compare the proportions of visitations in the *No Indicator* dataset to the *One Indicator* dataset. We used one-tailed tests due to our hypothesis of having higher proportions of visits to sites with privacy indicators. The results of these tests are depicted in Figure 7.5 and Table 7.4. Using Fisher's Exact tests, we found that privacy indicators did have an impact on visitations, significantly increasing the visitation rates to results further down on the search results page.

We explored the role of privacy indicators on visitation rates further by grouping all of the ranked search results together. We excluded the first search result on each page because many of these were navigational, and therefore users were likely to visit them regardless of whether or not they were annotated with privacy indicators. We performed a chi-square test and found that users were significantly more likely to visit search results beyond the first result when the additional results were annotated with privacy indicators, (8.67% (4,076) (No indicator) vs 12.42% (138) (One indicator), $p < 0.0001$, $\chi^2 = 19.13$). Thus, we found that *Hypothesis 2* is supported by our data: users are more likely to visit search results which are lower on the search results page if those results are annotated with privacy indicators.

7.4 Discussion

Based on this research, we find that when privacy indicators are annotated to search results, they do have a significant impact on which websites users choose to visit, especially when a website is annotated with a high-privacy indicator. This indicates that websites, in general, may be able to leverage the quality of their privacy protections to drive more traffic to specific sites.

[5]Searches that contained P3P errors were filtered out of this dataset.

	No Indicator	One Indicator	Fisher's Exact p
Result 1	61.16% (3,225)	57.69% (45)	0.30
Result 2	19.21% (1,015)	28.36% (19)	0.05
Result 3	12.81% (674)	25.00% (22)	**0.002**
Result 4	11.01% (575)	19.84% (25)	**0.003**
Result 5	7.96% (416)	13.11% (16)	0.04
Result 6	6.79% (355)	8.94% (11)	0.22
Result 7	5.18% (271)	8.85% (10)	0.07
Result 8	4.34% (226)	7.25% (10)	0.08
Result 9	4.96% (257)	4.91% (8)	0.58
Result 10	5.55% (287)	9.94% (17)	0.02

Table 7.4: Visitation rates for sets of search results when none of the search results had a privacy indicator and when exactly one result had a privacy indicator. Fisher's exact test was used to compare the proportions of visitations using the Bonferroni correction to account for multiple testing ($\alpha = 0.005$).

While field study data supported our two hypotheses, certain limitations in the experimental design should be kept in mind. While we asked users to use Privacy Finder as their normal search engine, it is sometimes hard to convince users to switch from an existing search engine, especially as search engines become more tailored to each individual user (e.g. Google's web history). Additionally, due to our use of a daily raffle ticket incentive, some users may have participated with a minimal amount of effort, performing one query, and then simply closing the browser. For example, we had ten cases where people who participated in the study for longer than one day conducted small numbers of searches for the days they participated (e.g. a user who conducts 6 searches over 5 days). However, since we eliminated sets of search queries where none of the results were visited, this may have cut back on confounding effects where participants were not actually interested in finding information. Additionally, we did not have a set of perfect control data for this study. Instead, we used the search result visitation patterns gleaned from the use of a large commercial search engine to validate the search visitation patterns of our data, and a within-study control of search result sets returned without privacy indicators. A preferrable situation would be to form a research partnership with a large search engine company. With this partnership, we would be able to work with the large search engine company to integrate Privacy Finder into their search engine, and deploy a larger scale Privacy Finder field study to a subset of their users for a specific amount of time.

In addition to privacy indicators, defaults have a strong impact on user interface settings. Our data analysis focused on searches made in the Google and Yahoo! search engines. While the Yahoo! Shopping search engine option was available for our participants, we did not collect enough data to specifically examine differences in general information seeking searches versus shopping based searches. The majority of searches (91%) were also conducted at the default privacy preference level (medium), suggesting that the privacy settings may have lacked real meaning to users.

Examining the impact of privacy information, we found that people are drawn to the high-privacy indicators. In our previous laboratory studies, we found that it was not the indicator (green boxes) itself that was the draw, but what those indicators symbolized [129, 49], in this case, privacy. Further work is needed to validate the impact of random indicators versus meaningful ones in the field to determine if people are

attracted to indicators, regardless of their meaning. Additionally, sites with P3P policies are relatively rare in the search results, and seeing the P3P indicators may be somewhat of a novelty. While we did not see a "novelty" effect in the search result visitations for the users who only used Privacy Finder for one day, the browsing patterns for larger samples of users may be different. This leaves an open question of the impact of P3P indicators once adoption rates have increased. Likewise, our dataset was too small to significantly examine the impact of multiple P3P results on a single page. This would be an interesting question to examine once P3P adoption rates increase.

Another avenue of research is that of the impact of privacy signaling. Privacy indicators may also be viewed as a proxy for reliable websites. Further research should be conducted into the extent that people take privacy indicators into account compared to other factors such as the design of the website or the brand name of the website.

7.5 Conclusions

People use Internet search engines to satisfy the majority of their informational needs. However, even though people are more concerned about their online privacy, they do not take the time to thoroughly examine the privacy policy of every website they encounter. The P3P standard was created to make this privacy information more accessible. Often, it is this lack of access to privacy policy information, or information asymmetry [3], that causes people to not act according to their privacy preferences. Thus, making privacy policy information available in the search engine can be a significant boon to users.

The results of this field study support our previous findings that people will seek out or visit sites with visible privacy ratings. Accessible privacy information does have an impact on search result browsing behavior. We find that the Privacy Finder search engine interface can act as an asset to both users and to websites that post P3P information. Users can choose to visit sites that better match their privacy practices. Websites can increase their visitation rates if they have P3P policies that search engines interpret and use as the basis for privacy indicators. The results of this study suggest that the adoption of P3P and the increased transparency for privacy policies will not have a detrimental effect on search result visitations, even if a website's specific policy may not be as good of a match to a user's privacy settings. Specifically, it can drive more clicks if the site is rated with a high privacy rating.

Chapter 8

Design Patterns

Design patterns provide software developers with generalized solutions to common recurring problems [24]. The first design patterns were intended to provide developers with reusable code to solve many common architectural problems [59]. However, as security and privacy problems have become more prevalent, several authors have published works on design patterns to improve privacy and security [108, 72, 152, 17].

C-HIP Stage	**Common Problems**	**Design Pattern(s)**
Attention Switch & Maintenance *Do users notice the indicator?*	Indicator was not prominent User was not looking for indicator User was not looking for absence	*Active Warnings* *Noticeable Contextual Indicators* *The Absence of Indicators*
Comprehension/Memory *Do users know what it means?* *Do users know what it wants them to do?*	Indicator was not read Indicator was confusing No choices/recommendations	*Providing Recommendations* *Attractive Options* *Conveying Threats & Consequences*
Attitudes & Beliefs *Do users believe the indicator?*	Previous experiences Environmental stimuli Indicator does not convey trust	*Levels of Severity* *Separating Content* *Immediate Options*
Motivation *Are users motivated to take recommended action?*	Something outweighs risk User does not understand risk User does not believe risk is relevant	*Separating Content* *Attractive Options* *Conveying Threats & Consequences*
Behavior *Do they do it?*	User does not know how to act User incorrectly acted	*Failing Safely*

Table 8.1: This table depicts the design patterns that I created to prevent common errors in the C-HIP model. The first column lists the stages of the C-HIP model, the second column gives examples of common problems, and the third column lists the appropriate design patterns.

Based on results of the studies I presented in the previous chapters, I created a set of design patterns for overcoming many of the errors that I discovered users make when viewing trust indicators. Table 8.1 shows an overview of these design patterns and how they relate to the various stages of the C-HIP model. Because of the many design considerations that go into trust indicators, the design patterns fit into a hierarchy of specificity ranging from whether to even show an indicator all the way to the intricate details of the indicator's layout. Each pattern is modeled to follow Tidwell's format [126]: the problem and *what* the

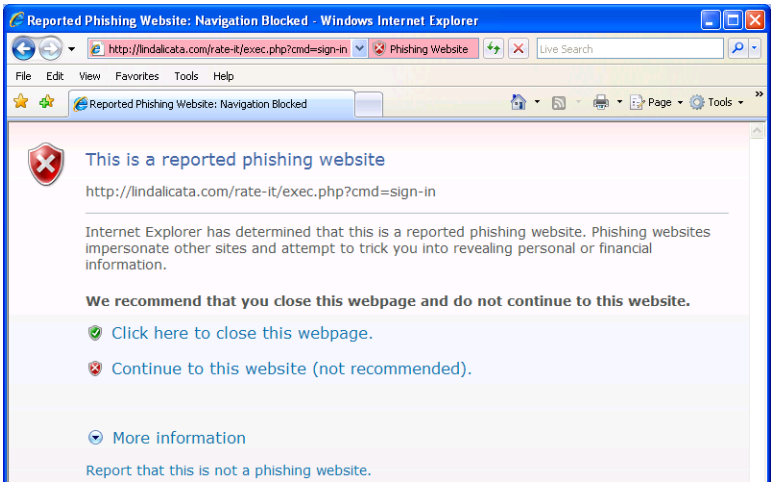


Figure 8.1: The active warning used by Internet Explorer 7.

pattern does to solve it, *when* it should be used, and *how* it can be implemented. I take this format a step further by also explaining how each pattern was motivated by the user studies that I performed, whether each pattern has any tradeoffs, and how each pattern may be used by an attacker.

8.1 Active Warnings

8.1.1 The Problem and Solution

Some warnings fail in very critical situations because they were not prominent enough for the user to notice them. "Active" warnings should be used to grab users' attention by interrupting their primary tasks, thus forcing them to acknowledge the warnings by taking an action in order to proceed.

8.1.2 When

Active warnings should be used when there is a significant reason to believe that the user is in imminent danger (i.e. ignoring the warning may result in adverse consequences such as falling for a phishing attack). Such warnings should only be used when suggesting a new course of action, not merely to provide contextual information. A warning is considered "active" when it forces the user to make a decision, likewise a "passive" warning either does not force interaction or can be dismissed without forcing the user to make a decision.

8.1.3 Why

These warnings are designed to force the user to take notice and increase the likelihood that the user will take the correct action. Passive warning styles that do not interrupt the user may go unnoticed and thus be rendered useless. Likewise, a warning may be passive if it can be dismissed without the user taking notice of it. Interrupting the primary task forces the user to complete the "Attention Switch" phase of the C-HIP

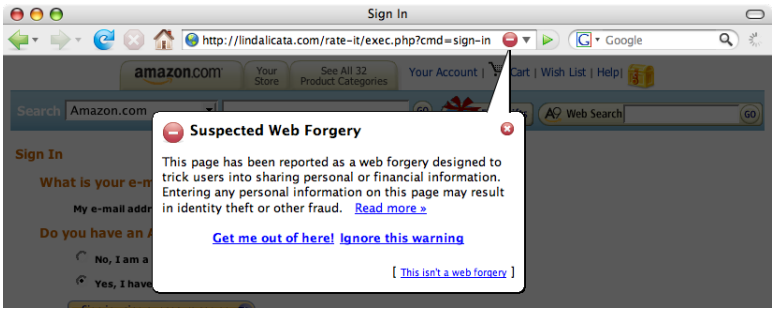


Figure 8.2: The active warning used by Firefox 2.

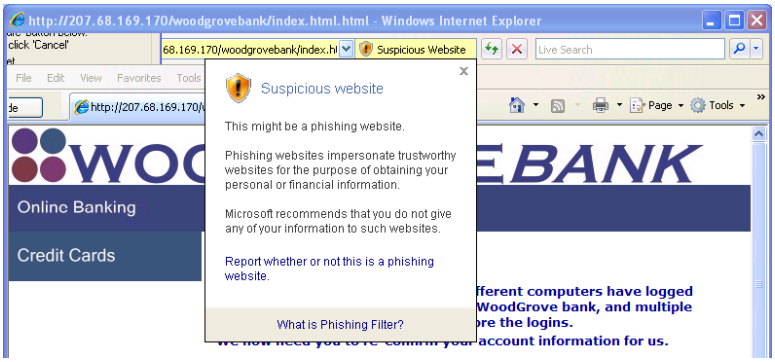


Figure 8.3: The passive warning used by Internet Explorer 7. This warning does not force user interaction; if a user clicks elsewhere in the browser window, the warning disappears.

model. This means that the common errors of users not explicitly looking for warnings or not noticing them can be minimized. By forcing the user to make a decision, errors at the "Attention Maintenance" phase of the C-HIP model are minimized because the user is forced to interact with the warning.

8.1.4 How

Active warnings must be designed to interrupt the primary task by either replacing the content users were expecting with the warning message, or by drawing attention away from the expected content. These techniques can be seen with the IE7 and FF2 examples (Figures 8.1 and 8.2): the website content is replaced with the full screen warning or the website content is dimmed and the warning is superimposed upon it. A poorly designed passive warning can be seen in Figure 8.11, where the user is not provided with any options and the warning can be dismissed by clicking anywhere on the webpage.

8.1.5 Motivation

In Chapter 3 I presented a study where we found that warnings that did not interrupt the users went unnoticed. In that study we found that in many cases using a passive warning was not significantly different than simply not providing a warning altogether. The passive IE7 warning took several seconds before appearing, during that time participants were already focused on entering their personal information. Their keystrokes on the

website caused the warning to be dismissed without them noticing it. This happened to two participants on both tasks such that these participants never realized that the warnings were ever displayed. This was not the case for the active IE7 and Firefox warnings: every participant in these conditions noticed the warnings. By interrupting users' primary tasks and forcing them to make a decision, significantly more users paid attention to the warnings and were ultimately protected from the phishing attack.

8.1.6 Considerations

This pattern describes a method of alerting users to an impending danger and therefore fits at the top of the trust indicator hierarchy; this pattern does not give strict guidance for the layout of a trust indicator, only the type of indicator that should appear given a high-risk situation. *Active Warnings* are generalizable to many other areas where there are threats that have the same risk level as phishing attacks or certain SSL errors (i.e. the domains I examined in this thesis). Risk level should be determined based on both the likelihood and severity of the danger. For instance, based on the number of websites that use expired certificates, the chance of encountering a malicious expired certificate is relatively low even though the consequences may be severe (e.g. identity theft, economic losses, etc.). Thus, *Active Warnings* should not be used to alert users to expired certificates because the risk is small. On the other hand, *Active Warnings* should be used to alert users to serious situations such as malware being detected or having unsaved documents open on a low-battery laptop. Risk level must seriously be considered before deciding to use an active warning. If the danger is unlikely or the consequences are minor or many users may not care about the consequences, users may become habituated to ignoring the warnings.

8.1.7 Subversion

The main way of subverting this design pattern is for an attacker to habituate a user into dismissing active warnings without reading them. This way, when a user encounters an active warning in a high-risk situation, she will be more likely to ignore the warning. One way of doing this could be through spam campaigns where HTML-based emails sent from botnets are designed similarly to current warning messages. Another way would be for attackers to create pop-up windows on websites that appear similar to warning messages. In both cases, users would become habituated to dismissing windows that look like active warnings. However, this may not be a serious threat if the *Failing Safely* design pattern is used in conjunction with an active warning.

8.2 Noticeable Contextual Indicators

8.2.1 The Problem and Solution

Contextual indicators are often not taken into account because users do not notice them. To prevent this, place the contextual indicators near the user's decision area or locus of attention.

8.2.2 When

Contextual indicators are used when additional information may help the user make a better decision. Thus, the indicators should be available at the time that the user makes that decision.

8.2.3 Why

This design pattern helps the user overcome common errors in the "Attention Switch" and "Attention Maintenance" stages of the C-HIP model. By placing the indicator near the user's point of focus, the chance that the user does not notice the indicator is minimized. Showing the indicators before the user is confronted with a decision or after a decision has already been made will force the user to make the decision without the aid of the indicator.

8.2.4 How

When a decision is being made within the web browser, the user's focus will be on the available options. These options should be annotated with additional contextual information. For instance, when presenting search results, icons can be used to annotate the search results (Figure 8.4). Thus, when the user is deciding which website to visit, he or she will notice the indicators next to the available choices.

8.2.5 Motivation

In the study performed by Wu et al., few participants noticed the indicator in the browser chrome because they forgot to look at a specific area of the screen before making a decision [146]. Whalen et al. performed a similar study where they noticed that few users noticed the SSL icon without first being prompted [135]. More recently, Sobey et al. found that the Extended Validation SSL indicators used by Firefox 3 were ineffective for the same reason [116].

In Chapter 6 I presented a study where we found that search results annotated with privacy indicators helped privacy-conscientious users find high-privacy websites. In one of the other conditions, the privacy indicators were presented to the users as frames above the destination websites, and therefore were not located in a place that users would normally be looking. When participants purchased batteries from the first website they visited, they paid significantly more money when they saw the privacy indicators as search result annotations ($t_{20} = 2.792$, $p < 0.011$), and rated privacy as a significantly greater factor ($t_{20} = 3.001$, $p < 0.007$), than those who saw the indicators above the destination websites. Likewise, when participants purchased sex toys from the first website they visited, they also paid significantly more money when they saw the privacy indicators as search result annotations ($t_{19} = 2.772$, $p < 0.012$). In the post-study survey, we found that all participants had similar privacy concerns, regardless of the experimental condition to which they were assigned. Thus, these differences in behaviors can be attributed to the placement of the privacy indicators: when they were displayed at the locus of attention—alongside search results—participants were more likely to take the indicators into account when deciding which websites to visit and subsequently patronize.

I validated this finding with the study I presented in Chapter 7. We had participants use Privacy Finder in the field for their everyday online queries. Search results with P3P policies were annotated with privacy indicators, much like they were in the *search* condition in Chapter 6. Not including the first search result of the pages, we found that search results that were annotated with privacy indicators were significantly more likely to be clicked than search results that were not annotated with privacy indicators ($p < 0.0001$, $\chi^2 = 19.13$). This confirms that the indicators were noticed when they were placed at the locus of attention.

8.2.6 Considerations

Just like the *Active Warnings* pattern, this pattern is very high-level and only describes where contextual indicators should be placed on the screen, and says nothing about how they should be designed. This pattern is also generalizable to many other areas beyond online privacy: any time contextual information needs to conveyed to users without interrupting their tasks. In the case of privacy, not everyone cares about privacy and the consequences of ignoring privacy may not be as severe as other online threats. Thus, *Noticeable Indicators* should be used when there is not a clear danger, but when some users may benefit from having additional information. Thus, such indicators may also be used to convey price information, merchant information, or even information about handicap accessibility. Likewise, because there is not a clear universal danger, the risks of users becoming habituated to these indicators are minimal.

8.2.7 Subversion

In theory, indicators placed in browser chrome cannot be altered by website content and are therefore trusted. However, these indicators often go unnoticed. The *Noticeable Indicators* pattern prevents this by placing indicators in locations where users are likely to be paying attention. An attacker may subvert the effectiveness of these indicators by creating spoofed versions on websites where they might not normally appear. For instance, an indicator representing merchant reliability may be copied and placed on the website of an unscrupulous merchant. This may confuse users and dilute the real purpose of the indicator. One way of preventing this is by limiting the indicators to third parties, such that the indicator appears on a website not owned by the website the indicator represents. For instance, a search engine may use contextual indicators to annotate websites so that the indicators are displayed before the user visits one of the websites represented by the inidcators. The third party (e.g. the search engine) could also trademark the indicators so that it would be illegal for someone to use them without permission. While these precautions would not stop a motivated attacker, these indicators are not designed to be a defense against serious security threats.

8.2.8 The Absence of Indicators

The Problem

Some indicators indicate positive things, thus, users are supposed to be alerted when these indicators do not appear on a webpage. However, many users are unalarmed by the absence of these indicators or simply fail to take notice.

The Solution

Users should not be expected to notice the absence of an indicator. Instead of using positive indicators when at a “good” website, use a negative indicator to indicate a “bad” website.

When

An indicator should be used when a danger has been detected, or additional information is available which may lead the user to believe the a website is unsafe.

Why

Building on *Noticeable Contextual Indicators*, absent indicators are rarely noticed because there is nothing to examine at the user's locus of attention. When positive indicators are used for trust, attackers can also mimic them and confuse the user. Most users do not know the difference between chrome and content, therefore there is an incentive for the attacker to spoof positive indicators. When negative indicators are used, there are fewer incentives for the attacker to spoof them (e.g. an attacker has a lower incentive to spoof a phishing warning message than a "Secured by VeriSign" logo). Users often do not remember to look for symbols or icons that are not pervasive. Therefore, these symbols should not be relied upon.

How

When using warning message to distinguish between good and bad websites, do not use indicators to denote good websites. Either use warning message when encountering bad websites, since there will be less of an incentive to spoof these, or use pervasive contextual indicators. Both types of indicators should be inserted by the web browser so that website designers will have a harder time spoofing the indicators (i.e. if the web browser will always add such an indicator, and a malicious website tries to spoof it, the user will likely see two conflicting indicators, which should raise suspicion).

Figure 8.5 depicts the SiteKey indicator. This trust indicator is likely to fail because users are expected to be alerted to a spoofed website when the indicator is absent. Another common indicator is the SSL lock icon, where users are supposed to be on alert when submitting sensitive information when this icon is not present.

Motivation

Whalen et al. noticed that users do not look for the presence of SSL icons on "good" websites, and therefore it is unlikely that users will notice their absence on "bad" websites [135]. Schechter et al. found that when removing the SiteKey indicator and replacing it with a generic message, almost every user still tried to log in, potentially compromising their credentials. They also found that when removing the SSL indicators, no users noticed [110].

Jackson et al. found that most users are unable to distinguish web browser chrome from website content. Thus, users are likely to fall for "picture-in-picture" attacks [76]. Adelsbach et al. also found that many web browsers are susceptible to various exploits that may allow an attacker to spoof SSL iconography [6]. Because of this, positive indicators will likely be spoofed whenever they are used by legitimate websites.

8.2.9 Considerations

The *Absence of Indicators* design pattern is a very high-level pattern because it describes the types of indicators that should be used for security situations, but does not specify the details of how these indicators should be designed. This design pattern only applies to indicators representing security, trust, and/or privacy, and may not be generalizable to other types of indicators. Specifically, this pattern is unlikely generalizable to other types of indicators when there is no incentive to spoof those indicators. This pattern may also prevent habituation by limiting the number of indicators to which users are exposed; there are likely more "good" websites than "bad" websites.

8.2.10 Subversion

Positive indicators, denoting "good" websites, are much less safe because they can easily be spoofed and users often do not notice when these indicators are missing. Negative indicators, denoting "bad" websites, can also be spoofed, but there is less of an incentive to do so. Instead, an attacker could subvert negative indicators by displaying them frequently in an attempt to habituate users. At the same time, users are no worse off then if positive indicators were used, since these indicators are much easier to spoof.

8.3 Providing Recommendations

8.3.1 The Problem and Solution

Many warnings fail, not because users did not understand what the dangers were, but because the warnings did not present clear suggestions on how to avoid those dangers. To prevent this, the warning message must provide the user with a suggested course of action and instructions on how to pursue that course of action.

8.3.2 When

When a danger has been detected, the user should be presented with a clear recommendation on how to safely proceed, as well as a list of other possible actions.

8.3.3 Why

If a warning highlights a potential danger and conveys the danger to the user, the warning may still fail if the user does not understand how to mitigate the danger. The recommendation must be present to explain *how* to proceed.

In many cases, users will read the title of a warning message and then skip to the available options. If no options are available, the user will likely make an uninformed decision (e.g. ignoring the warning because no recommendations have been presented). This design pattern addresses errors in the "Comprehension/Memory" stage of the C-HIP model that stem from the user not understanding what the warning wants them to do.

8.3.4 How

The recommended action should be more prominent than all other options. Thus, if a user elects to disregard the detailed description found in the body text and skip to the options, it will be easy to understand what the recommended action is. The available options should be designed such that it is trivial to distinguish the recommended option from the other available options.

Figure 8.6 depicts two warnings: the warning on the left is from IE6 and describes a danger but does not actually make any recommendations on how to proceed; the user is simply left with the choice to dismiss the warning. The warning on the right is from IE8. This improved warning displays the recommended option annotated with a green icon and larger than the alternate option. The alternate option, which is not recommended, is annotated with a red icon.

8.3.5 Motivation

Downs et al. performed a study on phishing and found that the "higher recognition than recall of warnings is typical of familiar but poorly understood stimuli." That is, many of the security warnings were ignored by users not because they went unnoticed, but because the users did not understand what they were supposed to do after seeing them [44]. Stoll et al. made a similar observation after performing a study on two graphical security systems: participants in the control condition made poorer decisions because they were unsure of what to do with the information they were provided [120].

In Chapter 3 I presented a study on web browser phishing warnings where I found that users of the passive warning in IE7, which provided no actionable recommendations besides dismissing the warning, were no better off than those who did not see any warnings. Several of these participants understood that the warning was saying something about a suspicious website, but they did not understand what action they were supposed to take in order to heed the warning. Thus, in the absence of recommended actions, they ignored the warning proceeded. Likewise, in Chapter 5, I presented a study on SSL warnings. Our two custom warnings and the IE7 SSL warning all provided recommendations, while the FF2 and FF3 warnings did not provide recommendations. When asked what they believed the warnings wanted them to do, the users of FF2 and FF3 were significantly less likely to understand that they should not submit personal information ($p < 0.0246$; 37.5% vs. 61.7%). Thus, providing a recommendation improves warning comprehension and helps users to understand what actions they should take to mitigate a given risk.

8.3.6 Considerations

Unlike the previous design patterns, this pattern is much more specific: *Providing Recommendations* provides guidance on how information should be displayed on a warning, and is therefore a much lower-level design pattern. All high-risk warnings must provide actionable recommendations for how a user is to proceed, otherwise the user is forced to simply dismiss the warning. This is generalizable to all warnings where there is a high risk of danger, but the system cannot automatically determine the most appropriate action to take, instead only offering a suggestion. If the system knows the proper action to take, it should automatically take that action and not bother the user. If the system does not know which action should be taken because it might vary from user to user, this pattern also does not apply.

Habituation is unlikely to be a problem for this design pattern because it does not specify how the recommendation should appear, only that it should be prominent. In general, users are unlikely to become habituated to the recommendation, but instead the warning as a whole, which is a concern for some of the other design patterns.

8.3.7 Subversion

It is not apparent how an attacker might subvert this particular design pattern, other than by convincing users that either there is no real danger, or by convincing her that the recommended option will not help her complete her primary task.

8.4 Attractive Options

8.4.1 The Problem and Solution

Users may read the options presented to them by a critical warning, but may not choose the recommended option because they either did not understand the threat or did not believe the recommended option would help them complete their primary task. In critical warnings, a recommended option may not be selected if the user does not believe it will allow her to complete a primary task. This error may be prevented by creating recommended options that appear conducive to completing the primary task. Additionally, labels on warning options should underscore the threat model so that if the user does not read anything else, she still understands the danger of ignoring the warning.

8.4.2 When

When displaying critical warnings, if multiple options are presented to the user, the recommended option should allow the user to complete the primary task (Figure 8.7). If only one option is presented (e.g. an acknowledgement), this becomes unnecessary because the user does not need to choose among several options. However, this acknowledgement should still be worded to underscore the threat model.

8.4.3 Why

If a user does not think that the recommended option will allow the completion of a primary task, she will not be motivated to obey the warning. Additionally, if the user does not read any other parts of the warning, she will still need to read the options in order to dismiss the warning. Therefore, the options should underscore the threat so that there is additional motivation to take the recommended option. This design pattern addresses errors in both the "Comprehension/Memory" and "Motivation" stages of the C-HIP model.

8.4.4 How

For critical warnings that contain multiple options, the recommended option should use wording that appears conducive to completing the primary task. This wording should also underscore the threat that the warning is attempting to guard against. For instance, a warning on a suspected malware website might say "search for an uninfected version of this program."

8.4.5 Motivation

In Chapter 4 I presented results from a study on phishing warning option text. Phishing relies on tricking users into visiting fraudulent websites that appear similar to trusted websites. The main way of distinguishing a phishing website from the legitimate one that it is spoofing is by examining the URL. In our laboratory study, we compared the recommended option to "go to my homepage instead" with a recommended option to "search for the real website." We found that when the option text emphasized that participants were visiting a fraudulent website (i.e. "search for the real website"), they were significantly less likely to trust the URL ($F_{2,42} = 4.469$, $p < 0.017$). That is, those who saw the former text were less likely to be suspicious of

the URL because the option text did not underscore the threat model. This indicates that carefully selected option text can prevent errors in the "Comprehension/Memory" stage of the C-HIP model.

If the recommended option does not appear to facilitate the primary task, users may perform a riskier option, which will likely cause them harm. In the case of phishing, which we examined in Chapter 4, this means visiting a fraudulent website. Upon initially viewing the warnings, participants in the *search* condition were 225% less likely to initially dismiss the warning by taking the riskier option: on the first viewing of the warning, 13% of participants took the riskier option in the *search* condition, whereas the riskier option was chosen by 30% of the participants in the two other conditions, where the recommended option was to "go to my homepage instead." Thus, participants were initially more willing to choose the recommended option when it appeared to help them complete their task of visiting a particular website. However, several of these participants viewed the warning again so that they could choose the riskier option after they were unable to find the real website, which ultimately caused them to fall for the attack. Overall, based on the option text, participants viewed the warnings significantly longer, which indicates that the wording of option text has the potential to prevent errors at the "Motivation" stage of the C-HIP model.

8.4.6 Considerations

Like the previous design pattern, *Attractive Options* is a low-level pattern because it specifies how text should appear on a warning, rather than a general design or guidelines for when to warn. Like all of the previous patterns, it should also be fairly generalizable; in any context if users are given two options, they will be more likely to choose the option that seems likely to help them. However, if the user chooses the recommended option and later regrets that choice, she may be unwilling to choose that option in the future. In this manner users may become habituated to choosing one particular option if the same option text becomes sufficiently pervasive. This could be prevented by making the warning text more dynamic such that the options change based on the actual circumstances in which the warning is being displayed.

8.4.7 Subversion

This design pattern may be subverted if attackers can successfully habituate users to the warnings. For instance, the IE warnings use a green shield icon to indicate the recommended option. An attacker could use this same iconography in other contexts in an attempt to dilute its meaning. Another way that this design pattern could be subverted is if an attacker can convince the user that the "attractive" option is not really that attractive. For instance, if the text of a phishing warning recommends that the user "search for the real website," and the attacker is spoofing a website that does not really exist, the user may tire of searching and reluctantly choose the option to proceed despite the warning. These attacks rely on confusing the user and may not be easily countered.

8.5 Conveying Threats & Consequences

8.5.1 The Problem and Solution

Users may ignore the indicator because they do not believe it applies to them. To prevent this, the indicator should succinctly convey the threat it is representing as well as the potential consequences of ignoring it.

8.5.2 When

In the case of critical warnings, the description text in the warning should at a minimum explain why the user is seeing the warning and what the possible consequences of ignoring the warning are.

8.5.3 Why

If users notice the indicator, but do not understand why it is appearing, they may be unwilling to follow the indicator's suggestions. For instance, if the warning uses jargon to describe a threat so that the user does not understand it, or the warning simply fails to describe the threat at all, the user is likely to ignore the indicator. This design pattern address problems in the "Comprehension/Memory" stage of the C-HIP model.

Likewise, if users notice the indicator, understand the indicator, understand the actions that the indicator wants them to take, and believe the indicator, they still may not take those actions because they may not believe that the consequences apply to them. For instance, a phishing warning may be ignored if a user incorrectly thinks she is protected by her anti-virus software. Therefore, this design pattern also addresses errors in the "Motivation" stage of the C-HIP model.

8.5.4 How

The wording to describe threat details, consequences, and how to mitigate those consequences should be written succinctly without using jargon. This text should appear between the heading and options of the warning to increase the likelihood that it will be read (Figure 8.12).

8.5.5 Motivation

In Chapter 4 I presented results from a study on phishing warnings. In that study we found that most participants who ignored the warnings and were phished did so because they did not believe the warnings applied to them. These warnings did not explicitly mention the threat model and therefore many users incorrectly believed that they were not in any danger because they were using a computer that did not belong to them. In fact, when asked about the danger of ignoring the warnings, only 24% of the participants correctly mentioned the theft of their personal information or someone else inappropriately accessing their accounts. Participants in the *search* condition initially made better choices and spent significantly longer analyzing the warnings ($F_{2,41} = 4.75$, $p < 0.014$) because the improved option text helped to convey the threat. However, because they did not understand the possible consequences for ignoring the warning, these participants were ultimately just as likely to make the same poor choices as those in the other study conditions.

In Chapter 5 I presented results from a study on SSL warnings where we redesigned the warnings to emphasize risk. The redesigned warnings stated the threat model, the consequences of ignoring the warning, and how to mitigate those consequences. We tested this warning alongside the existing FF2, FF3, and IE7 SSL warnings. We discovered that when viewing our new warnings, participants were more likely to understand the threat model, the consequences of ignoring the warnings, as well as the actions that the warnings wanted them to perform. Thus, by succinctly providing this information to users, users will be more motivated to act based on increased risk perceptions. This prevents errors at both the "Comprehension/Memory" and "Motivation" stages of the C-HIP model.

8.5.6 Considerations

The *Conveying Threats & Consequences* design pattern is a low-level pattern that describes the text that should appear in a critical security warning. Based on all of my work on both phishing and SSL errors, I have found that risk perceptions are one of the largest motivators for users' decisions of whether or not to obey a security warning. Because of this, this design pattern should be generalizable to any other type of security warning so that users have the opportunity to understand risks. At the same time, this is the most likely design pattern to succumb to habituation problems: the descriptive text of a warning is usually the first thing to get ignored. Since this design pattern only specifies blocks of text, there is little recourse once a user has become habituated. Once habituated, the user is likely to ignore the text when making a decision. This may be minimized by using the *Levels of Severity* design pattern.

8.5.7 Subversion

There does not appear to be a clear way of subverting this design pattern, since an attacker has little incentive to state the consequences of failing for an attack. However, an attacker could bombard the user with similarly-design warnings in an attempt to habituate her to ignoring the warning text, similar to the attack on the *Active Warnings* design pattern.

8.6 Levels of Severity

8.6.1 The Problem and Solution

Habituation occurs when similar-looking warnings are used for varying threat levels. Thresholds should be drawn for threat levels, such that warnings for differing threat levels should be distinguishable from each other.

8.6.2 When

When a risk is detected and a warning is presented to the user, the system should determine the relative risk level when deciding how the warning is to be displayed.

8.6.3 Why

If a user encounters a particular warning during a relatively low-risk situation, she may choose to disregard this warning due to the low risk level. If she encounters a warning that looks very similar during a high-risk situation, she may disregard this warning because it was confused with the low-risk situation. Figure 8.9 shows two SSL warnings from Firefox 2. These warnings address two different threats of different severity, but use similar designs, which may be confusing to many users.

This design pattern addresses errors in the "Attitudes/Beliefs" stage of the C-HIP model. Users may have prior beliefs about a particular type of warning, and thus may confuse similar-looking warnings with the current warning. Prior experiences with less severe warnings should not cause users to be habituated to critical warnings.

8.6.4 How

Warnings should be designed based on their risk level. This risk level should be determined based on the likelihood of the danger, the damage that may be caused to the user by ignoring the warning, and the likelihood that the warning may be triggered in error (due to a false positive). Too many differing discrete risk levels may overwhelm users, causing them to suffer from "warning overload." However, too few risk levels may result in habituation, causing users to ignore many critical warnings because they were confused with less-critical warnings. Future research is needed to determine the balance between habituating and overwhelming the user.

8.6.5 Motivation

Amer and Maris found that many users are habituated to the most common Windows warning messages because these warnings all have the same design. They found that habituation occurs after only seeing a warning a few times [9]. Brustoloni and Villamarín-Salomón found that habituation could be prevented by dynamically creating the text of warning messages [19]. However, it is unclear what the long term implications are for this method (i.e. when the user may become overwhelmed).

In Chapter 3 we found that the IE7 phishing warnings were ignored because users confused them with similarly designed SSL warnings that they had become habituated to dismissing. IE7 and FF2 were released at relatively the same time and included redesigned phishing warnings. Participants were assigned to these two web browsers in our laboratory and were asked whether they recognized the phishing warnings. We discovered that significantly more participants recognized the IE7 warnings than the FF2 warnings ($p < 0.048$ for a one-tailed Fisher's exact test). Many of them mentioned that they had seen these warnings while visiting internal work-related websites as well as university websites. Since they were visiting trusted websites, they said that they thought it was safe to proceed when they saw this warning. As we predicted, these participants were confusing the phishing warning with the less-severe SSL warnings that they had already become habituated to dismissing. Thus, their preconceived attitudes and beliefs about the IE7 warnings were adversely impacting their decisions in our study.

In Chapter 4 I showed that this problem can be minimized by adding a red border to the warning, thereby making it appear more severe and more distinguishable from warnings that represent lower risk levels. Participants who saw a red border around the phishing warning were half as likely to confuse it with existing IE warnings. Likewise, participants who saw the red borders spent significantly longer viewing the warnings ($F_{2,41} = 4.75$, $p < 0.014$), which indicates that they did not have pre-existing beliefs about them, and therefore had not already made up their minds about how to react before viewing the warnings.

Once a warning is redesigned based on its level of severity, care should be taken to ensure that it is only displayed when a matching hazard is encountered. In Chapter 5 I showed that by showing SSL warnings only in high-risk situations, users are more likely to heed the warnings. We used a question about the destination website to determine whether the user would be likely to enter personal information at a suspicious website. Based on this question, we determined whether or not the participant was actually at risk, and accordingly determine whether or not to display a critical warning. Participants who were not at risk saw significantly fewer warnings and therefore may be less likely to become habituated to the warning representing this particular risk level. The SSL warnings used by FF2, FF3, and IE7 do not change based on estimated risk levels. Thus, participants who saw FF2 and IE7 SSL warnings were likely to ignore the warnings on both websites, regardless of the underlying risk levels. This may be because they had become habituated to

dismissing these warnings on websites that did not pose a threat, and therefore dismissed the same warnings when there was a much greater risk of danger because of their preconceived notions. These types of errors occur at the "Attitudes & Beliefs" stage of the C-HIP model and can be prevented by this design pattern.

8.6.6 Considerations

The *Levels of Severity* design pattern is a very high-level pattern, just like *Active Warnings* and *Noticeable Contextual Indicators*, because it describes the overall look and feel of the warning, as opposed to specific details about a warning's contents. Habituation is a serious problem for warnings, and may be a forgone conclusion given enough exposures to similarly-designed warnings. This pattern can slow habituation by making warnings distinguishable based on their severity. With too few warning designs, because of too few levels of severity, habituation may occur quicker than if there are more levels of severity. However, too many levels of severity—corresponding to too many different warning designs—may result in warning overload and user confusion. Future studies may need to be conducted to determine the optimal number of severity levels and corresponding warning designs. Likewise, once users become habituated to a particular warning design, that warning design should be changed. Ostensibly this can be done during software version updates.

8.6.7 Subversion

This particular design pattern might be subverted if an attacker intentionally tries to habituate users to a particular warning design, much like the attack described for the *Active Warnings* pattern. This could be mitigated by periodically changing the warning designs. More importantly, if warnings are rarely displayed, users will be less likely to become habituated. We can accomplish this by one of two ways: automating security decisions so that warnings are not needed, and by showing warnings only when absolutely necessary. Advances in detection technology would allow developers to better predict threats and automatically counter them. For instance, in the case of phishing, an ideal solution would be to detect that a user is trying to visit a phishing URL, and then redirect them to the correct website instead. This would obviate the need for a severe warning as the user would no longer be in danger, yet still allowed to perform the intended task.

8.7 Separating Content

8.7.1 The Problem and Solution

Users often let the "look and feel" of the website determine their level of trust, often to the detriment of unbiased trust indicators. Indicators should distort or not display the destination website such that the look and feel are not taken into account when the user is asked to make a trust decision.

8.7.2 When

When a critical warning is displayed, the website should be distorted or hidden. When contextual indicators are used to make trust decisions about which website to visit, the indicators should be presented before the user views the content of the website that was chosen.

8.7.3 Why

When a website is presented alongside a trust indicator (Figures 8.10 and 8.11), the user may use the look and feel of the website to determine the veracity of the trust indicator. Since many users are unaware of how easy it is to design professionally looking fraudulent websites, they may take the design quality into account when choosing to ignore a warning or contextual indicator (which may indicate that the website is not trustworthy).

This design pattern addresses errors in the "Attitudes/Beliefs" stage of the C-HIP model, which address both critical warnings and contextual indicators. This design pattern also applies to contextual indicators during the "Motivation" stage of the C-HIP model, because even if they believe a contextual indicator, the presence of website content may demotivate users from factoring the indicator into trust decisions. Environmental stimuli, such as the look of the destination website, should not detract from the amount of trust conveyed by an unbiased indicator.

8.7.4 How

In cases where a critical warning is about to be displayed, distort or hide the original website such that the user's focus is on the warning message. In cases where a contextual indicator is to be displayed, display it before the content of the website that it represents. For instance, this can be accomplished by annotating hyperlinks or by providing popups during mouse-overs.

8.7.5 Motivation

Fogg et al. have conducted several surveys which found that the "look and feel" of a website is often the most important factor when a user chooses to trust the website [55]. According to Egger, "the more a company is perceived to have invested in its web site, the less likely it is perceived to act opportunistically by betraying customers' trust [50]." This is why phishing is so effective: Dhamija et al. conducted a study of phishing websites and observed that 23% of the participants only used the content to determine a website's veracity. They concluded that current security indicators fail a large percentage of users because website design is a larger part of their trust decisions [42].

In Chapter 3 I presented a study of web browser phishing warnings. We found that participants who viewed IE7's passive warnings tended to not trust the warnings because they were shown alongside website content. While many participants saw the warnings, they distrusted the warnings because the website content looked credible, despite the fact that the website content was spoofed to look like a credible website. Specifically, three of the nine IE7 users who ignored the passive warning said that they did so because they believed that the destination websites looked authentic, and they therefore did not believe the warning when it told them that the website was suspicious. Thus, the website content caused them to err during the "Attitudes & Beliefs" stage of the C-HIP model by distrusting the warnings. We did not observe this effect when the warnings were shown instead of the website content—the active IE7 condition—or when the warnings distorted the website content—the FF2 condition.

In Chapter 6 I presented a study of different ways of displaying contextual indicators that represented website privacy policies. All of the study participants were concerned with privacy, but when viewing privacy indicators alongside website content, participants did not always make the same privacy decisions as those who did not see privacy indicators alongside website content. Participants in the *frame* and *interstitial*

conditions differed based on whether they saw the privacy indicators alongside website content or on a separate page before viewing website content. When participants saw the indicators before the website content (the *interstitial* condition), they were twice as likely to report factoring the indicators into their purchase decisions (26% vs. 50% of participants). This shows that participants who saw indicators alongside website content (in the *frame* condition) erred during the "Motivation" stage of the C-HIP model because they were less motivated to factor the privacy indicators into their decisions.

We also observed potential errors at the "Attitudes & Beliefs" stage of the C-HIP model, with regard to contextual indicators. During the study in Chapter 6, when we asked participants how much they cared about privacy and whether their privacy concerns were a factor for their purchase decisions, we observed no significant differences between the conditions. Because privacy was just as much of a factor regardless of whether or not they saw website content alongside the privacy indicators, and because participants made better privacy-protective decisions when they did not see website content alongside the privacy indicators, it is likely that the website content caused some participants to not believe the privacy indicators. This would indicate errors at the "Attitudes & Beliefs" stage of the C-HIP model, which may also be prevented by this design pattern.

8.7.6 Considerations

The *Separating Content* design pattern is a high-level pattern because it specifies when critical warnings and contextual indicators should be displayed, and does not guide the design or content of these indicators. In the case of critical warnings, it is not clear whether this pattern applies to areas outside of online security. When warning about online security threats, the veracity of a given website is called into question and therefore runs the risk of diluting from the warning if it is displayed alongside the warning. In other non-security critical warning use cases, the warning might not be competing with the hazard for the user's attention. For instance, take the case of a user on a laptop that is low on power. If a warning is used to prompt the user to save an open document, the open document is unlikely to cause the user to question the accuracy of the warning, and therefore this design pattern may not be appropriate. In fact, there may even be an argument in favor of displaying this type of warning alongside the document. Likewise, this pattern may also not be generalizable for contextual indicators. An indicator claiming a website is untrustworthy might not be trusted when displayed alongside the website. However, an indicator representing shipping time or total price might not be adversely impacted by being displayed alongside website content.

This pattern is less likely to have problems with habituation than other patterns. Content displayed alongside the indicators might increase the chances that the indicators are ignored, especially if the content is distracting and the indicators have become commonplace. By displaying the indicators apart from distracting content, there is an increased chance that users pay attention to them because there are fewer factors to consider.

8.7.7 Subversion

There are several ways in which an attacker may subvert this design pattern. In the case of contextual indicators, an attacker might force a fraudulent indicator to be displayed alongside the destination website's content, after the correct indicator was displayed before the user saw the content. For instance, Privacy Finder might annotate a search result as having zero green boxes, but the website may fraudulently post a similar indicator with four green boxes. In this case, users might trust the latter indicator more. In the case

of critical warnings, an attacker might use a script or exploit a browser vulnerability to pop up a window alongside a warning. Any content in this window might detract from the message of the warning. In both of these cases, it is unclear how these attacks could be mitigated.

8.8 Immediate Options

8.8.1 The Problem and Solution

Users may choose to ignore a warning because it presents an easy way of dismissing it. Instead, the option to dismiss the warning should not be immediately displayed.

8.8.2 When

In the case of critical warnings, when the warning first appears, the means for dismissing the warning should not be the most prominent feature.

8.8.3 Why

If users notice the indicator, understand what it is saying, but see that it is easy to dismiss, they may not believe that it represents a serious threat. This design pattern address problems in the "Attitudes & Beliefs" stage of the C-HIP model.

8.8.4 How

The option of dismissing a warning and not following the recommended option should not be the most prominent feature. Instead, either make the recommended option more prominent, require the user to take several steps to dismiss the warning, or hide the means of dismissing the warning altogether.

8.8.5 Motivation

In Chapter 3 I presented results from a study on phishing warnings. In that study we found that several participants said that they did not believe the warning were very serious because they were given the option to proceed anyway. One participant commented, "since it gave me the option of still proceeding to the website, I figured it couldn't be that bad." In Chapter 4 I presented a followup study on phishing warnings, where we made similar observations: several users thought that if the warning were serious, it would not make it easy for them to ignore it.

In Chapter 5 I presented results from a study on SSL warnings where we redesigned the warnings to emphasize risk. We tested our new warnings alongside the FF2, FF3, and IE7 warnings. Both the IE7 and FF2 warnings allowed the user to dismiss them with a single click, whereas the FF3 warning made it much more difficult, and our custom warnings obscured this option by using a tiny font away from the user's locus of attention. Both the IE7 and FF2 warnings performed significantly worse because the option to dismiss them was immediately obvious. Wu found similar results in his thesis work: when phishing prevention mechanisms give users an easy way of dismissing the warning and proceeding, they will likely do so [145].

8.8.6 Considerations

The *Immediate Options* design pattern is a fairly low-level design pattern because it guides both the design and content of critical warnings. This pattern is likely generalizable to many other areas, as there is a lot of anecdotal evidence that users tend to swat away dialog boxes and other types of warnings when it is easy to do so [95]. This pattern attempts to address a habituation effect, but it may also suffer from habituation: once a user does figure out how to override the warning, she may do so and become habituated to overriding it if she does not perceive a risk. Thus, this design pattern may only create a learning curve and delay habituation, but not prevent it.

8.8.7 Subversion

The easiest way for an attacker to subvert this design pattern is by helping users learn how to dismiss the warnings. For instance, in the case FF3, it took several steps to dismiss the warnings. Most users simply obeyed the warnings because they could not figure out how to dismiss them, even though they may have wanted to dismiss them. Attackers could circulate detailed instructions under the guise of "computer help" that will teach users how to easily dismiss the warnings in hopes of habituating people.

8.9 Failing Safely

8.9.1 The Problem and Solution

If a user does not attempt to comprehend a warning and instead opts to take whatever action he or she believes will simply dismiss the warning, the warning will not serve its purpose. This can be prevented by designing warnings such that if the user does not comprehend the recommended action, the warning performs a safe default action. Likewise, the default option should be the most prominent one so that it is obvious what the user should do.

8.9.2 When

The recommended action should be made apparent as soon as the warning loads, and before the user is expected to read any of the other warning text.

8.9.3 Why

If a user has no interest in learning why a warning was presented, he or she may attempt to execute the quickest action to dismiss the warning, thereby continuing the primary task. If the recommended action appears to the user as the quickest way of responding to the warning, he or she will likely take this action.

This design pattern addresses errors in the "Behavior" stage of the C-HIP model. A user may understand the risk, understand the warning and what it recommends, but still fail because they did not know how to take the recommended action. This pattern helps by increasing the likelihood that the user takes the recommended action. In this case, it is likely that more effort will need to be expended to take a riskier action.

8.9.4 How

The recommended action should be the most prominent of all the possible choices presented to the user (Figure 8.13). This can occur through the use of colors (e.g. the recommended action is colored green, whereas all the others are colored black or red), text size (e.g. the recommended action appears bigger than the other choices), affordances (e.g. clicking a familiar icon results in the recommended action), etc.

8.9.5 Motivation

Carroll and Carrithers found that users made fewer errors using a word processor when they were unable to interact with complicated features [21]. Whitten and Tygar built on this work by examining user errors in security software and found that many users made mistakes when using the PGP software because the most secure actions were not the most obvious ones [137]. They proposed the concept of *safe staging* where complicated features are not enabled by default until the user is capable of understanding them [136]. Xia and Brustoloni applied this theory to web browser security warnings after concluding that current warnings fail because they "do not tell users how they might overcome security errors: the software simply asks user permission to continue a task [147]." That is, security software fails when the safest option is not the most prominent one.

In Chapter 3 I presented a study on phishing warnings where we observed a few users who did not read the warnings but were still protected. These participants saw the phishing warnings, did not read them, and so the simplest actions that these users knew to perform was to close the browser window. They repeated this process until finally giving up and moving on to the next task. While the warning failed in that it did not alert them to an impending danger, they were still protected because they performed the recommended action inadvertently: they were forced to close the browser window because that was the most obvious action. We saw similar behavior among participants in the SSL study described in Chapter 5.

8.9.6 Considerations

The *Failing Safely* design pattern is neither high-level nor low-level because it vaguely describes a function that all critical warnings should include, without specifying the details of how that feature should be implemented. Every warning, regardless of what it represents, should fail safely so that users are compelled to make the best decision regardless of whether or not they understood or even read the warning. If a user takes the default action and finds that it was unhelpful, she may return to the warning and choose another—likely riskier—course of action. In this manner, she may become habituated to performing an unsafe action even when seeing a similar-looking warning in different contexts. The *Levels of Severity* pattern may mitigate this effect.

8.9.7 Subversion

If users become habituated to taking the default action because of a familiar affordance or other characteristic of the warning, an attacker could potentially exploit this by using these affordances in other situations. If users are making decisions out of habit and not attention to risk, there is little that designers can do to prevent this, other than periodically redesigning warnings to counter habituation effects. Additionally, designers can

use the *Conveying Threats & Consequences* pattern to increase the chances that users will make decisions based on risk.

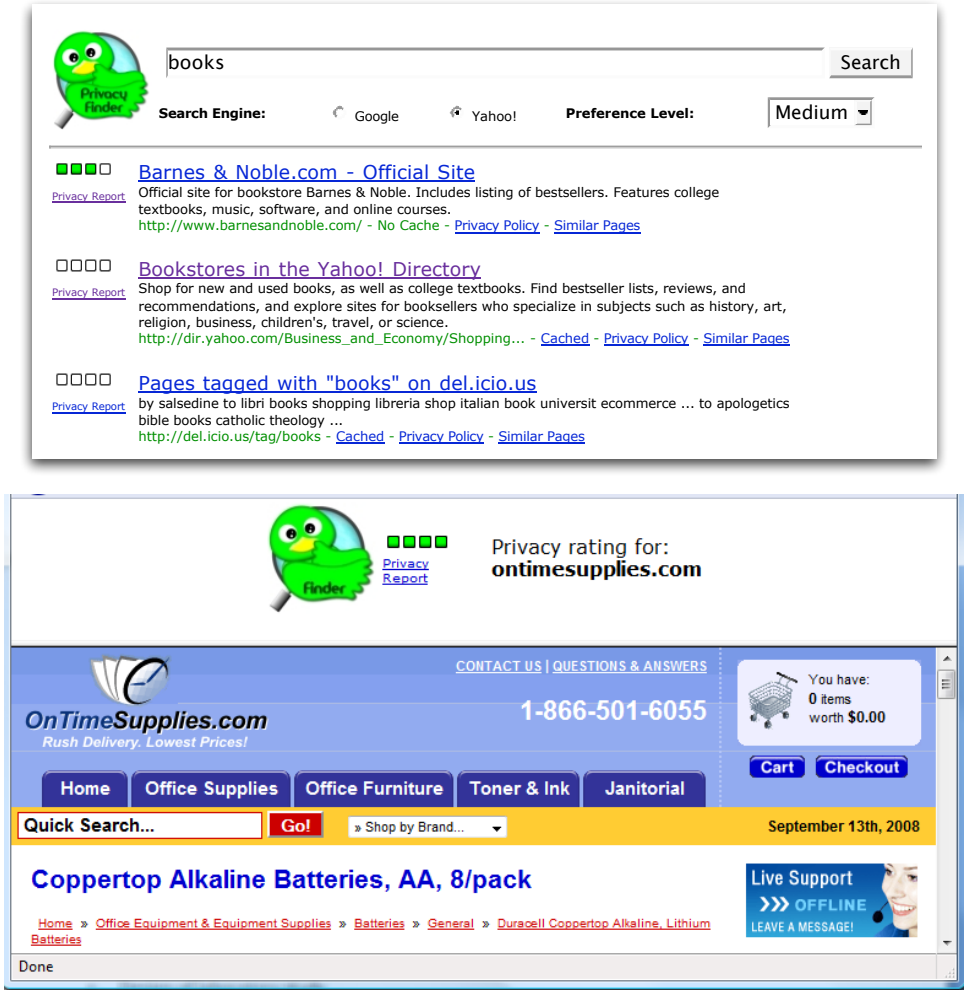


Figure 8.4: The contextual indicators used by Privacy Finder. These indicators are placed next to each search result where the user is likely to be looking (above). Thus, the user will be more likely to take the indicators into account when choosing a search result. We found that when placing the indicators above the destination websites (below), the indicators were less effective.

Figure 8.5: The SiteKey indicator as used by PNC bank. For this security indicator to be effective, the user is required to notice the absence of the tiger picture on a spoofed PNC website.

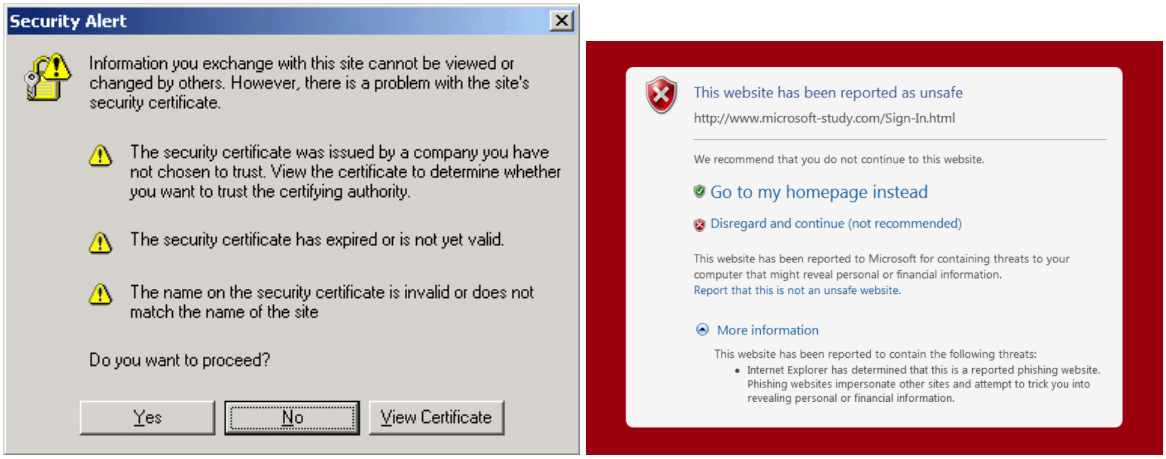


Figure 8.6: The warning on the left, from IE6, appears when a problem was encountered with an SSL certificate. The warning does not give the user any recommendation on how to proceed. The warning on the right, from IE8, appears when a user visits a suspected phishing website. The recommended option is annotated with a green icon and is larger than the option that is not recommended.

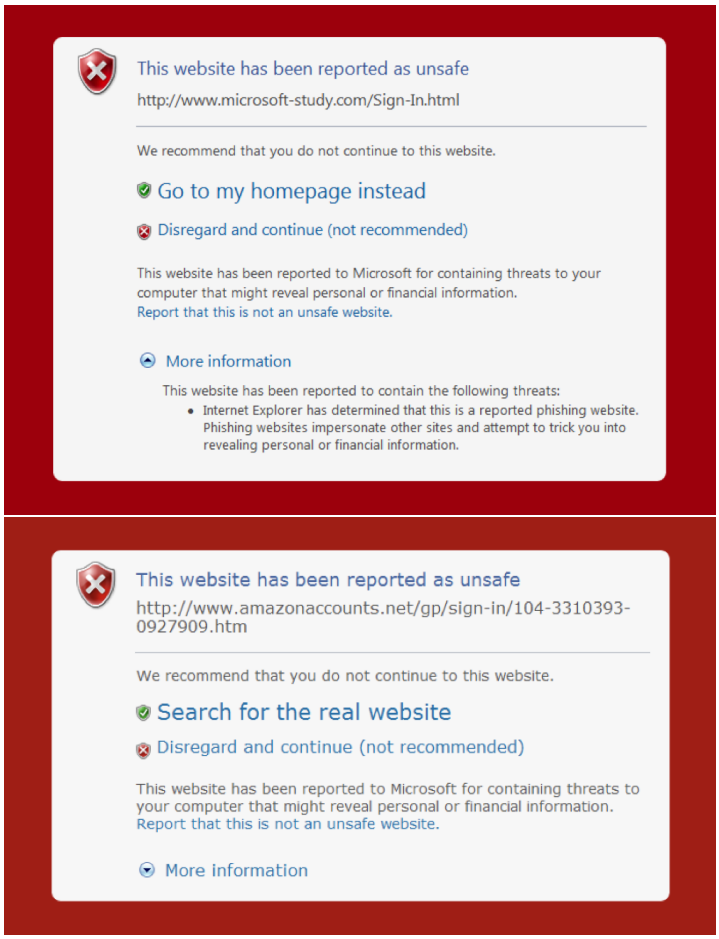


Figure 8.7: The top phishing warning recommends users "go to my homepage instead," which does not facilitate the primary task, nor does it underscore threat model. The bottom phishing warning recommends that users "search for the real website." This text facilitates completion of the primary task by helping the user locate the website she was initially trying to visit, as well as underscoring the threat model: she is currently visiting a fraudulent website.

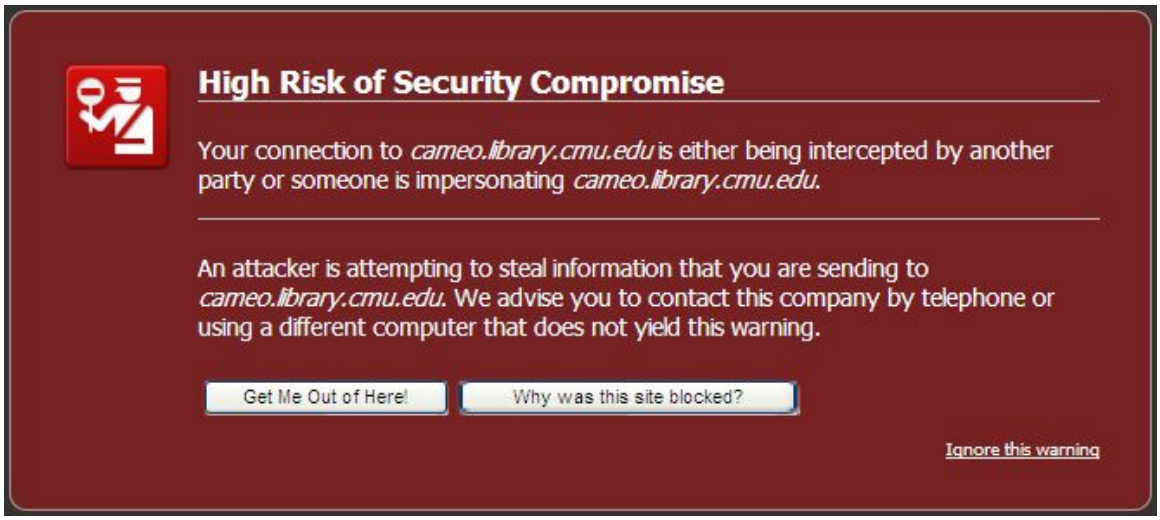


Figure 8.8: This newly designed SSL warning clearly states the threat it is guarding against, the consequences of ignoring it, and how to mitigate the risks.

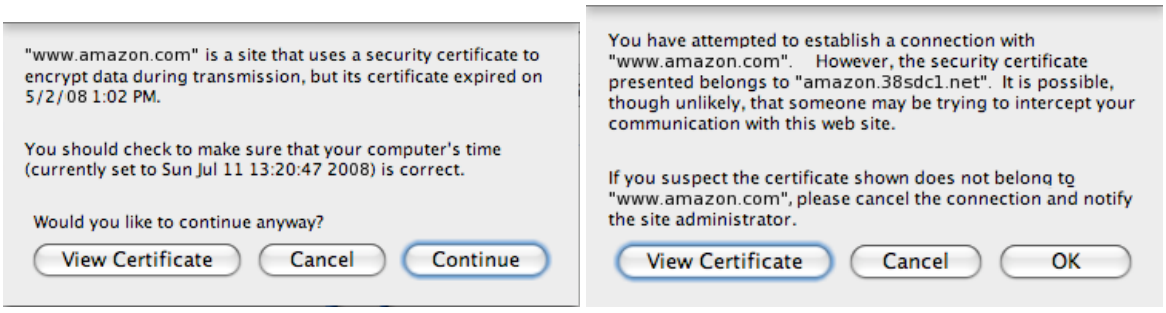


Figure 8.9: These two SSL warnings appear in Firefox 2 when the user encounters an expired certificate (left) or a certificate for a different domain name (right). Arguably the latter is a much more serious security threat, though both warnings are designed very similarly, and therefore may not be readily distinguishable.

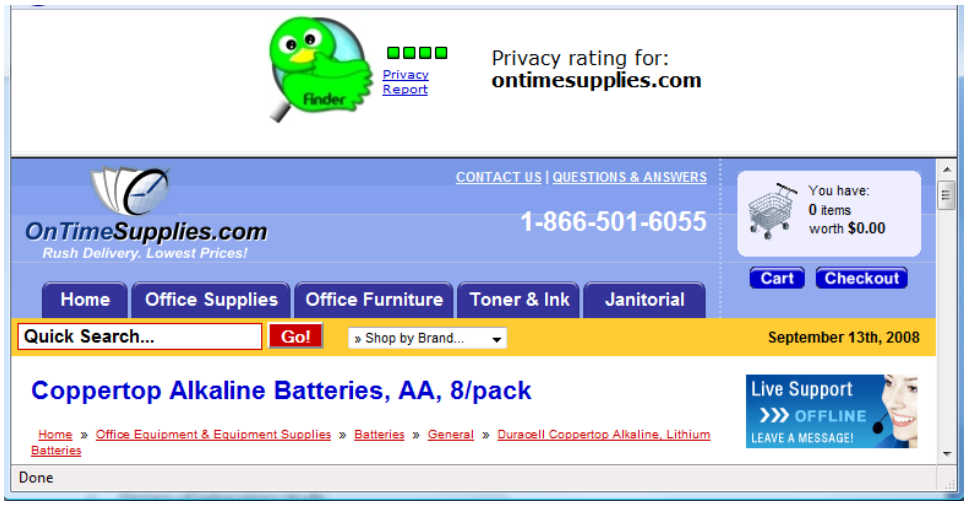


Figure 8.10: This privacy indicator appears above the content of the website such that the user is allowed to weigh the "look and feel" of the website alongside the privacy indicator. If a user is captivated by a website's content, it may cause the user to weigh the indicator less in her trust decisions, or even worse, she may incorrectly believe the indicator is in error.

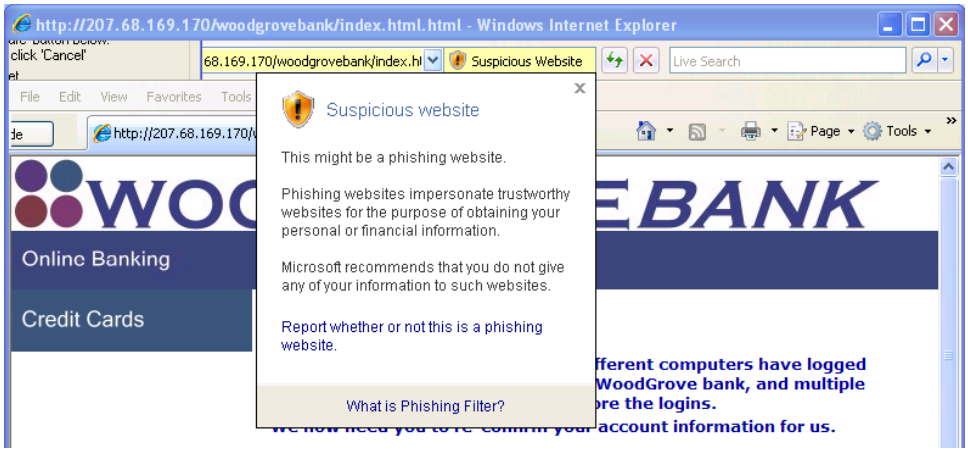


Figure 8.11: The passive warning used by Internet Explorer 7. This warning appears alongside the website content and may cause the users to trust the content more than the warning.

Figure 8.12: This new SSL warning presents the unsafe option, "ignore this warning," in very small text and away from the user's locus of attention so that it is not immediately obvious how to dismiss the warning.

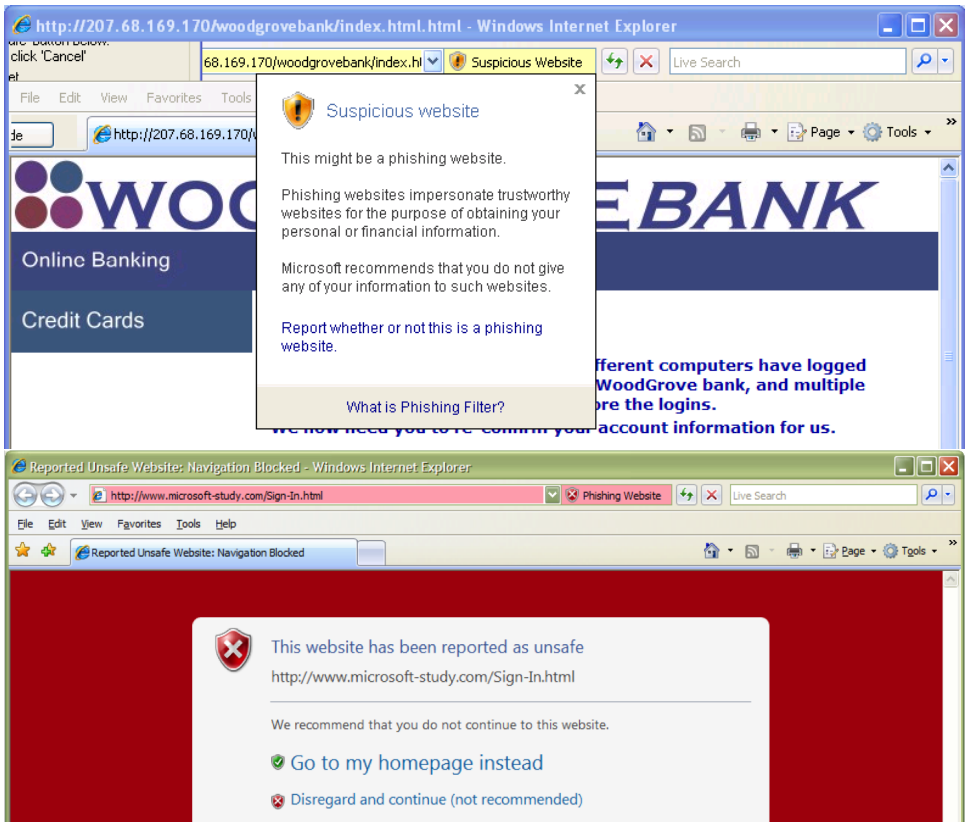


Figure 8.13: The passive warning used by IE7 (top) does not make it easy to perform the recommended action because the only obvious option is to dismiss the warning. The active IE8 phishing warning (bottom) solves this problem by making the recommended option appear more prominent than the riskier alternative. Additionally, if the user does not read the warning, the most obvious action is to close the window, which results in a safe action.

Chapter 9

Conclusion

In this thesis I presented the results of five user studies that I performed. Three of these studies were performed on *critical warnings*, while two of these were performed on *contextual indicators*. Based on my findings from these studies, I created a set of design patterns that I presented in Chapter 8. In this chapter I compare and contrast these design patterns with previous design patterns and show how they address common errors at each of the various stages of the C-HIP model. Specifically, I show how many errors can be minimized when designing critical warnings and how these design decisions differ for contextual indicators.

9.1 Previous Patterns

Yoder and Barcalow created some of the first design patterns to solve security problems. Their patterns addressed problems such as providing users with secure authentication mechanisms, privilege separation, and application security [152]. Since then, many people have authored design patterns for security, usability, and privacy. However, only a few patterns exist that align all three.

Van Duyne et al. created a set of design patterns for designing websites. Included in these patterns are several patterns that relate to both online security and privacy [132]. Specifically, their "Privacy Policy" pattern recommends that websites post conspicuous privacy policies. Romanosky et al. specify a similar pattern, which they call "Informed Consent for Web-Based Transaction" [108]. However, we know from the literature cited in Chapter 2 that most users do not bother to read natural language privacy policies. This was the motivation for my "Noticeable Contextual Indicators" pattern. Van Duyne et al. also created a pattern to address phishing. Their "Preventing Phishing Scams" pattern attempts to address the problem of phishing by recommending user education [132]. While education may be effective, security warnings should also be used as a last line of defense, as I recommend in my "Active Warnings" pattern.

In his thesis, Garfinkel proposed a set of design patterns for usable security, which have been further analyzed by Ferreira et al. [53]. Most of Garfinkel's patterns pertained to encryption and secure deletion of files, however several of his more general patterns are relevant to the ones I proposed in Chapter 8 [61]. Specifically, his "Create a Security Lexicon" pattern recommends that jargon should be defined in one central location. While this is reasonable advice for designers who are forced to use jargon, the results of the user studies I presented in this thesis on critical security warnings indicate users may be unwilling to

look up the jargon at the time that they encounter the warnings. Instead, I recommended in the "Conveying Threats & Consequences" that jargon should not be used at all so that the threat can be easily conveyed to as many users as possible without having them resort to reference materials. Unlike most other security design patterns that I have encountered, Garfinkel does provide patterns for security warnings. His "Warn When Unsafe" pattern specifies situations when designers should trigger warnings, however it does not specify how those warnings should be designed.

9.2 Critical Warnings

Critical warnings used by web browsers are designed to protect users from an impending danger. Such warnings are a last line of defense against a potential danger and should only be used when there is a high risk of danger, the risk cannot be automatically mitigated, or there is a potential for false positives. In this thesis I conducted two studies on web browser phishing warnings (Chapters 3 and 4) to determine how these warnings could be improved to prevent various types of user errors. Based on these findings, I designed an improved SSL warning and conducted a followup study to validate my improvements (Chapter 5). In this section I revisit these findings and the resulting design patterns in order to show how they prevent common errors in the stages of the C-HIP model.

9.2.1 Attention Switch & Maintenance

A warning may fail because users simply fail to notice it. When it goes unnoticed, users may never become aware of the impending danger that it is trying to warn them about. In Chapter 3, we saw that this was the case with IE7's passive phishing warning. During our laboratory study, we observed several participants who never saw the warnings because they were easily dismissed with keystrokes. To combat the problem of users accidentally dismissing security warnings without noticing them, I created the *Active Warnings* design pattern. I validated this by creating an active SSL warning. By interrupting the users' primary tasks and forcing them to make decisions about how to proceed, it was not possible for this warning to go unnoticed. In the laboratory study described in Chapter 5, we found that this warning was significantly more effective than the pop-up SSL warning used by FF2.

9.2.2 Comprehension/Memory

A warning may fail because users did not understand what it wanted them to do. Users may see the warning, but because they did not comprehend it, they may fail to take the correct action. In Chapter 3, we saw that there were several users of IE7's passive phishing warning who read the warning yet still decided to proceed to the phishing website. This was because this warning did not provide them with an actionable recommendation, and therefore they were unsure of what it wanted them to do. I created the *Providing Recommendations* design pattern in order to address these types of problems. I validated this design pattern by conducting the user study described in Chapter 5. When users were given clear recommendations on how to proceed, in the case of both our custom warnings and the IE7 SSL warning, they were more likely to understand what the warning wanted them to do.

9.2.3 Attitudes & Beliefs

A warning may fail because users do not believe what it is trying to tell them. Users may see the warning, understand it, understand what it wants them to do, but they may not heed it because they did not believe it was warning about a credible threat. This may happen because of habituation: if users see warnings in low-risk situations and become accustomed to dismissing them, they may dismiss all similarly designed warnings in the future, because they believe that all of them represent the same risk level. In Chapter 3 I presented results from a study on phishing warnings. We found that IE7 users were significantly more likely to dismiss the warnings because they had become habituated to dismissing similarly-designed SSL warnings. In order to minimize habituation, I created the *Levels of Severity* design pattern. By using different warning designs for different levels of severity, low-risk warnings are easily distinguishable from high-risk warnings. I validated this in Chapter 4 by adding a red border to the phishing warning, which made it easily distinguishable from the SSL warnings. We found that users of the red bordered warnings were less likely to recognize them.

Warnings may also fail when users simply do not believe them. Several web security threat models center around fraudulent websites masquerading as legitimate websites. If a phishing warning is shown alongside a website that is spoofing a legitimate website, users may base their decision of whether or not to trust the warning on the design quality of the website they are viewing. Thus, they may distrust the warning simply because they believe they are at the correct website. In Chapter 3 I presented results from a study on phishing warnings where we saw that several IE7 users made this mistake when they were viewing the passive warnings shown alongside the website content. I created the *Separating Content* design pattern to combat this problem; critical web browser security warnings should be displayed before website content is displayed so that users are not biased by that content.

9.2.4 Motivation

A warning may fail because users do not feel motivated to take the recommended action. Users may see the warning, understand it, understand what it wants them to do, believe the warning, but they may not heed it because they do not believe it applies to them. In Chapter 4 I presented results from my second study on phishing warnings. In that study we found that several participants decided to ignore the warnings because they mentioned generic security threats and thus users believed they only applied to malware. Since the users were not using their own computers, and therefore did not care about malware infections, they ignored the warnings and entered information into our phishing websites. Of course, believing that one is protected from all security threats by security software is not necessarily unreasonable. Internet users face a myriad of security threats; the consequences and possible attack vectors are very nuanced. Anti-virus software is likely the most common type of security software that Internet users install, but it does not protect them from many attack vectors other than malware, and many users do not seem to understand this. Therefore, it is understandable that many believe they are protected from most threats. At the same time, anti-virus software partially exists because people engage in risky behaviors (e.g. executing files of unknown provenance). It would be a Sisyphean task to teach users about every conceivable threat model and expect them to remember how to behave in any given situation. Instead, it might be easier to teach users about which of their behaviors may be risky in an online security context and how to avoid those risky behaviors. While security threats change over time, safe and unsafe behaviors stay relatively the same.

To counteract the problem of not understanding threats, I created the *Conveying Threats & Consequences*

design pattern. By clearly stating why the user is seeing a warning, as well as the possible consequences for ignoring that specific warning, users can make better decisions about whether or not the warning applies to them. I validated this design pattern in Chapter 5 by creating a custom SSL warning that clearly highlights the threats and consequences of accepting an unverified certificate. We found that when confronted with this warning, significantly more users understood the risks.

Warning options can also be used to motivate users to heed a warning's recommendation. The *Attractive Options* design pattern accomplishes this by using warning options to underscore a potential danger, as well as to help users complete their primary tasks. In Chapter 4 we created a phishing warning that had an option that said "search for the real website." This option emphasized that the users were likely visiting a spoofed website and helped them to find their intended website. We discovered that significantly fewer users incorrectly used the URL of the spoofed website as a factor in their decisions when they saw this option.

9.2.5 Behavior

A warning may fail because users do not understand how to take the correct action. Users may see the warning, understand it, understand what it wants them to do, believe the warning, feel motivated to take the recommended action, but it still may fail if they are unable to take that action. Don Norman called this problem the *Gulf of Execution* [96]. This problem can be minimized in security warnings by making the recommended option the easiest action for a user to take. For instance, in Chapters 3 and 5 we observed several users who did not bother to read the warnings and instead closed the web browsers. In this sense the warnings *failed safely* because despite the fact that they went unread, they forced the users to take the recommended action because that was also the simplest action.

9.3 Contextual Indicators

Contextual indicators can be used by web browsers to supply interested users with additional information with which they may use to make a decision. Such indicators have differing design concerns from those of critical warnings since the latter attempt to protect all users from an impending danger, while the former exist to supply only interested users with information. In this thesis I specifically looked at contextual indicators used to supply privacy information. In Chapter 6 I examined how best to display contextual indicators representing website privacy policies, while in Chapter 7 I validated these findings by performing a field study. In this section I review my design patterns for contextual indicators, I show how they prevent errors at the various stages of the C-HIP model, and I also compare them to design concerns for critical warnings that I discussed in the previous section.

9.3.1 Attention Switch & Maintenance

Just like critical security warnings, contextual indicators may fail users when the users simply fail to notice the indicators. In Chapter 6 I presented a study on various methods of displaying privacy indicators. We found that by displaying the indicators near a user's locus of attention—in this case, in the content pane of a web browser—the user is more likely to notice the indicators. I created the *Noticeable Contextual Indicators* design pattern to specify how such indicators can be displayed to minimize the chances of going unnoticed. I validated this design pattern in Chapter 7 by performing a field study to examine whether users would use

the indicators when performing web searches in their natural environments. We found that search results annotated with privacy indicators were significantly more likely to be selected, which indicates that these indicators were being noticed.

With regard to errors in the *Attention Switch & Maintenance* stages of the C-HIP model, contextual indicators have different design considerations than critical warnings. In the case of the former, the indicators are provided to give the user additional information about a website. With that information, the user can better form an opinion about a website. If the user does not notice the indicators, she does not find herself in a dangerous situation, she simply does not have all available information. Whereas critical warnings are the last line of defense against a very real threat. If a user does not notice a critical warnings, she is likely faced with a very serious risk.

9.3.2 Comprehension/Memory

Errors in the *Comprehension/Memory* stage of the C-HIP model can be detected by examining two different things: whether users understand the indicator and whether they understand what it wants them to do. In this thesis I examined contextual indicators that were labeled with text in order to convey their meaning. I did not control for these labels and so cannot make claims as to whether this is the best method of conveying meaning for contextual indicators. Contextual indicators have different constraints than critical security warnings, because the latter exist to guide users around a clear and present danger. Therefore, they make recommendations about how a user should proceed. Contextual indicators differ because they are not used to guard against a clear and present danger. They exist to provide information so that users can make more informed choices, but the indicators do not make recommendations because the risk they are guarding against is nuanced: not all users may care about the information that the contextual indicators are providing. Therefore, the *Providing Recommendations* design pattern does not apply to contextual indicators.

9.3.3 Attitudes & Beliefs

Contextual indicators may fail users when the users do not trust that the indicators are accurate. Wu et al. performed a study on contextual indicators that provided information about potential phishing websites and found that when the indicators were displayed alongside website content, users chose to distrust the indicators because they believed that the websites looked trustworthy [146]. I examined this phenomenon in Chapter 6 of this thesis by displaying privacy information above website content. While the data for this condition was not as significant as we had hoped, we still observed that users made different choices when the indicators were presented alongside website content as opposed to when they viewed the indicators before viewing the website content. Based on this, I created the *Separating Content* design pattern.

Contextual indicators and critical warnings share some of the same design decisions with regard to preventing errors at the *Attitudes & Beliefs* stage of the C-HIP model. Both types of indicators can be undermined by the presence of fraudulent-yet-well-designed website content. Designers should be careful to pay attention to how unverified content may interact with trust indicators. At the same time, the *Levels of Severity* design pattern may not apply to trust indicators since the indicators are purely informational and are not intended to guard users from imminent harm.

9.3.4 Motivation

Contextual indicators may fail when users ignore them because they do not believe they are relevant. In addition to the *Attitudes & Beliefs* stage of the C-HIP model, the *Separating Content* design pattern also prevents errors at the *Motivation* stage of the C-HIP model. If contextual indicators are displayed alongside website content, users may still believe the indicators, but the content of the website may captivate their attention so that they are less motivated to factor the indicators into their decisions. In Chapter 6 I presented a study on different types of privacy indicators. In one condition we showed indicators alongside website content, while in the others we showed the indicators before participants viewed the content. We found that participants were slightly less likely to base their decisions on the indicators when they saw the website content.

The *Conveying Threats & Consequences* design pattern also prevents errors at the *Motivation* stage of the C-HIP model. I did not test this pattern with regard to contextual indicators. In all the studies of contextual indicators, we allowed participants to click the privacy indicators to view additional information about each website's privacy policy. This information included the consequences of sharing personal information with the website, as per this design pattern. However, we did not collect enough data to validate this design pattern with regard to contextual indicators. Additionally, since screen real estate is limited, especially when dealing with contextual indicators, additional considerations may need to be made with regard to how this information should be displayed.

9.3.5 Behavior

Finally, contextual indicators differ from critical warnings because they do not make recommendations about a specific action and therefore do not suffer from errors at the *Behavior* stage of the C-HIP model. Therefore, the *Failing Safely* design pattern only applies to critical warnings.

9.4 Future Work

During the course of this thesis I performed five different user studies to create and validate the design patterns that I presented in Chapter 8. While all of these studies yielded new information about how people perceive trust indicators, and most of these studies yielded significant results regarding my hypotheses, there are several areas that require further inquiry. Specifically, I plan to conduct additional studies to examine the role of website content on trust indicators, the importance of selecting good option text for critical warnings, and to examine long-term effects of habituation.

9.4.1 The Role of Content

Fogg et al. conducted a series of surveys and determined that the "look and feel" of a website is usually the greatest factor in a user's trust decision [55]. Thus, when confronted with a contextual indicator that is displayed alongside the website content, users may choose to ignore the indicator because they believe the website "looks trustworthy." Unbeknownst to them, a professionally designed website says nothing about the policies and reputation of the website owner. In Chapter 6, I performed a study on privacy indicators and created an experimental condition in an attempt to control for website content. In this condition, the privacy indicators were displayed alongside the website content, whereas in the other experimental conditions they

were not. While we found that user behaviors significantly differed in a few cases when content was displayed to them, the effect was smaller than what we had expected. This may be because we did not control for the "look and feel" of the website content.

A future study could be conducted to examine the extent to which a well-designed website detracts from privacy indicators. I envision a study that is designed similarly to the one presented in Chapter 6, however, prior to the laboratory experiment, an online survey would need to be conducted. Differing online vendors would be used in this study, and the survey would contain screenshots of their websites. Survey respondents would rank each website using a Likert scale based on how trustworthy the site appears to them. Based on the aggregate scores for each website, we can then control for both price, privacy, and perceived "look and feel" of each website. This would then give us a better idea of how website content influences users' decision to trust privacy indicators.

9.4.2 Option Text

In Chapter 3 of this thesis I showed that users of IE7 disregard its phishing warnings far more frequently than they should. As a result, Stuart Schechter and I were given the opportunity to redesign the warnings in IE8. While our suggestions were taken seriously and a number of components of our design were adopted, they were selected and integrated in such a way that we suspect the improvement in user behavior will be far below what we had hoped. Specifically, we were interested in examining the extent to which the red background and choices for option text impacted user behaviors. We performed a laboratory experiment to study these effects, which I presented in Chapter 4. Unfortunately, the effect sizes were smaller than we expected, so we are planning to conduct a field study with a much larger sample size.

Our field study will consist of between six and eighteen hundred participants. Conditions will be randomly assigned such that there will be between one and three hundred participants in each condition. Participants will receive an email inviting them to participant in our study in exchange for an entry into a raffle for a gift card or other gratuity. To participate, participants will follow a link in the invitation email to a website hosted by us. On this website they must enter a valid Live ID (or sign up for a Live ID) to proceed. After entering a Live ID, participants will be presented with a software download. This software installs the new warnings in IE, randomly assigns them to an experimental condition, and instruments the warnings to report back to us.

We envision using six experimental conditions to control for the background color, the option text, and the descriptive text. These conditions are design to further test my *Levels of Severity*, *Attractive Options*, and *Conveying Threats & Consequences* design patterns.

Several days after downloading and installing our software, participants will receive an email from us explaining that we are offering a second raffle for another gift card if they visit a URL contained in the message and use their LiveID to sign up. We plan on using the same email that was described in Chapter 4. This message will be sent from a domain other than Microsoft.com (Microsoft-study.com), and the website where they sign up will also be from a domain other than Microsoft.com (Microsoft-study.com). That website will contain a form to enter their Windows Live ID (though no information will actually be transmitted). This website will also be on the phishing blacklist, which will cause the phishing warning to appear. Using javascript in the warning, we will report back to our servers whether the users ignored the warning and visited the website anyway. No Live IDs will actually be collected, but we will log whether or not users attempted to submit their information.

9.4.3 Habituation

When a user sees several similar-looking warning messages over a period of time, habituation can occur. Once habituated, the user may simply ignore a warning or confuse it with a similarly-designed—but different—warning. This may cause users to take unsafe actions. We saw that this was the case with the IE7 phishing warnings presented in Chapter 3. In this thesis I proposed creating different designs for security warnings based on their level of severity to prevent users from confusing a serious warning with a less-serious one. This was validated in Chapter 4 with the addition of a red border to the IE phishing warnings and in Chapter 5 with the creation of a new SSL warning. While in both cases our laboratory users were less likely to confuse these warnings with other less-serious warnings, it is unclear whether this effect will last over a period of time. There are many factors that contribute to habituation that need to be studied further through controlled laboratory and field studies:

- Frequency of exposure to a warning
- Time period of exposure to a warning
- Similarity to other warnings
- Consequences when previously ignoring/obeying warning
- Perceived consequences of ignoring/obeying current warning

This is not a comprehensive list and only serves to illustrate a handful of factors. I believe it is necessary to conduct preliminary studies to build a more comprehensive list of habituation factors. I would then plan on examining the extent to which each factor contributes to habituation, and how the different factors interact with each other. This research is of benefit to designers, engineers, and other researchers because it will guide the design of future security—and possibly other—warnings. I envision a future where security warnings are rare because the software is able to make decisions without user-intervention. Until then, effective warnings are still necessary.

Appendix A

Phishing Warning Study Recruitment Survey

Online Shopping Study - Recruitment Survey

Exit this survey >>

Thank you for your interest! This Carnegie Mellon University research study on online shopping will give you $35 to shop online. We expect you to buy two different items online. You should expect to keep around $20 as well as the items purchased.

You will receive the full payment on the day of the study.

*** 1. Are you still interested in participating in this study?**

- [x] Yes
- [ ] No

Next >>

Online Shopping Study - Recruitment Survey

Exit this survey >>

Please answer the following questions:

∗2. Indicate how often you use the following websites:

	Never Used	Use 1-10 Times/Year	Use 1-10 Times/Month	Use Daily
Amazon.com	○	○	○	○
eBay.com	○	○	○	○
PayPal.com	○	○	○	○
Banking Online (any bank)	○	○	○	○

∗3. Do you have an eBay.com account?

- ○ Yes
- ○ No

∗4. Do you have a PayPal.com account?

- ○ Yes
- ○ No

∗5. Have you purchased something online in the past year?

- ○ Yes
- ○ No

∗6. Can you check your email from someone else's computer (e.g. using a web browser)?

- ○ Yes
- ○ No
- ○ Don't Know

∗7. Are you currently using a Mac or a PC?

- ○ Mac

- PC
- Don't Know

8. Which web browser are you currently using?

- Internet Explorer
- Netscape
- Firefox
- Safari
- Opera
- Don't Know
- Other (please specify)

<< Prev Next >>

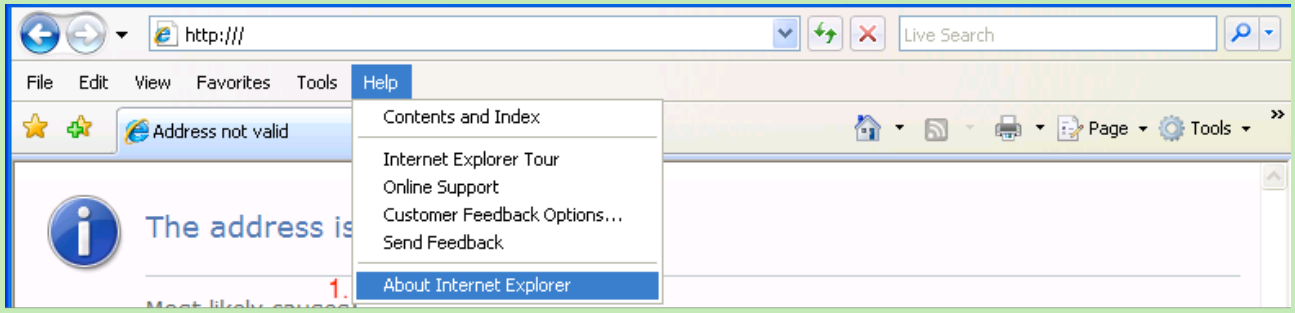
Online Shopping Study - Recruitment Survey

Exit this survey >>

On this page you will determine your browser version.

For users of Internet Explorer, to determine your browser version:

0. Open a new window using the menu: File -> New Window
1. Go to the "Help" menu and click "About Internet Explorer."
2. Locate the version number from the popup window.



∗10. What browser version are you using?

<< Prev Next >>

Online Shopping Study - Recruitment Survey

Exit this survey >>

These questions concern your experiences online.

***11. Have you ever participated in a research study at CMU before?**

- Yes
- No

12. If yes, what was the purpose of the study?

***13. Do you have an online store/vendor that you often visit or purchase from?**

- Yes
- No

14. If yes, what store(s) or vendor(s)?

15. Please enter whether or not you have been subjected to the following:

	Yes	No	Don't Know
Credit card fraud			
Stolen online password			
Stolen Social Security Number			
Identity theft			

<< Prev Next >>

Online Shopping Study - Recruitment Survey

Exit this survey >>

The contact information that you provide us will only be used for scheduling an appointment for participation in our study. We will not use this information for any other purpose.

***16. What is your name?**

***17. What is your email address?**

***18. What is your phone number?**

***19. Gender:**

- Female
- Male

***20. What is your age?**

21. What is your occupation?

***22. Have you ever...**

	Yes	No	I'm Not Sure
Designed a website?			
Registered a domain name?			
Used SSH?			
Configured a firewall?			

<< Prev Next >>

Appendix B

Phishing Warning Study Exit Survey

CMU Shopping Study Exit Survey

1. Browser Specific Questions

∗1. Which browser did you use in this study?

- Internet Explorer
- Firefox

∗2. Before this study, had you ever seen the specific warnings used in this study?

- Yes
- No
- I'm not sure

∗3. Did you read the full text of the warnings? Why/why not?

∗4. When the warnings were displayed to you, what was your first reaction?

∗5. What did you believe the warnings meant?

∗6. What action(s) did you think the warnings wanted you to take?

∗7. Did you believe the warnings?

8. How do you think the suspicious URL got to you?

***9. Please explain why you chose to either heed or ignore each of the warnings.**

***10. How much did the following factors influence your decision to heed or ignore the warnings?**

	No Influence at all: 0	1	2	3	4	5	Strongly Influence: 6
The text of the warning							
The colors of the warning							
The choices that the warning presented							
The destination URL							
The look and feel of the destination website							
Other factors (please describe below)							

11. If there were any other factors, please describe them.

***12. Which factor had the most influence on your decision?**

***13. How many of the warnings did you completely read?**

Neither of them

- One of them
- Both of them

14. For the warnings you read, why did you read them?

15. If you took heed of any of the warnings and chose not to visit the pages, why did you do so?

16. If you ignored any of the warnings, why did you ignore them?

Next >>

CMU Shopping Study Exit Survey

2. Online Habits

∗17. How much time do you spend on the Internet per week?

- 1 to 5 hours
- 6 to 10 hours
- 11 to 20 hours
- 21 to 30 hours
- More than 31 hours

∗18. How many email messages do you receive on average each day?

- Less than 10
- 10-30
- 30-50
- 50-100
- 100 or more

∗19. Can you describe what is meant by "phishing"?

<< Prev Next >>

CMU Shopping Study Exit Survey

3. Phishing Specific Questions

"Phishing" is when a con artist sends you a deceptive message claiming to be from someone else. The message will contain a URL to a website that will look very similar to a legitimate website, but if you enter any information, it gets sent to the con artist. This can result in accounts being compromised, credit card fraud, and identity theft.

*** 20. Have you ever received any phishing messages?**

- Yes
- No
- Don't Know

21. In a given week, how many phishing messages do you receive?

- None
- 1-5
- 5-10
- 10-20
- More than 20

*** 22. Do you know anyone who has ever entered personal information at a phishing site in the past (this does not include during the course of this study)?**

- Yes
- No

*** 23. Has your web browser ever warned you about suspected phishing sites in the past?**

- Yes
- No
- Don't Know

<< Prev Next >>

CMU Shopping Study Exit Survey

4. Online Security Questions

∗24. Have you ever had any online account information stolen?

- Yes
- No

∗25. Have you ever found fraudulent transactions on a bank statement?

- Yes
- No

∗26. Have you ever had your social security number stolen?

- Yes
- No

∗27. Have you ever been notified that your personal information has been stolen or compromised?

- Yes
- No

<< Prev Next >>

CMU Shopping Study Exit Survey

5. Demographics

∗28. What is your age?

∗29. What is your gender?

- Female
- Male

∗30. What is your highest level of education?

- Some high school
- High school diploma
- College degree
- Graduate Degree
- Professional degree (including trade school)
- Other (please specify)

31. How would you describe your race and ethnicity?

- White
- Black
- Asian or Pacific Islander
- Latino(a)/Hispanic
- Native American
- Other (please specify)

32. What is your country of origin?

33. If you have any additional comments, please write them below.

Thank you for completing this questionnaire! Please raise your hand to notify the study administrator to receive your $35 payment.

<< Prev Done >>

Appendix C

Warning Options Study Instruction Sheet

Instructions

We are studying how you interact with your email. We will be observing you with an eye tracker during the study. This study will last roughly an hour.

Please sit up as straight as possible in order to ensure proper function of our eye tracker. We may ask you to adjust your position during the study if you move too much.

- To ensure you receive several emails during the study, we will try to send you a message every ten minutes.

- When you read, act upon, or respond to an email, we cannot advise you in any way.
    - Act as you would outside of our laboratory.
    - Once you've begun to take action, the actions you take will not affect your prize.

- You are welcome to browse the web when not reading email.
    - Please keep Internet Explorer in full screen mode.

Prizes

- You have already qualified to receive a software gratuity for showing up.
    - You will receive the gratuity regardless of your performance.
- We will also send you an Amazon.com gift card, to which we will add:
        - $1 for every new email read
            - *New messages are those that arrive after you open your mailbox.*
        - $4 for every email that you take some action upon.
          For example:
            - Responding to a friend
            - Visiting a website
            - Forwarding a message
    - You will receive a maximum of $30.
    - You will not receive more than $5 per email message.
    - Any conversations or interactions initiated by you are ineligible for a reward.
    - Amazon.com gift cards will be sent within six weeks (though they usually arrive much faster).

Appendix D

Warning Options Study Exit Survey

Exit Survey

Participant#:_____

These questions pertain to the warning that IE displayed when you visited the Windows Live Challenge Website.

1. **Before this study, had you ever seen the warning that Internet Explorer displayed? (Circle one)**

   Yes No

2. **Did you read the full text of the warning? (Circle one)**

   Yes No

   Why/why not?

3. **When the warning was first displayed, what was your initial reaction?**

4. **How likely is it that something bad would happen if you continued on to the website after seeing this message? (Circle one)**

   0% 25% 50% 75% 100%

5. **What did the warning recommend that you do?**

6. **What do you believe are the possible consequences of disregarding this warning?**

7. **Please explain why you chose to either heed or disregard the warning.**

8. **How much did the following factors influence your decision to heed or ignore the warning?**

a.	**The text of the warning:**	0	1	2	3	4	5	6 (Strongest)
b.	**The warning color:**	0	1	2	3	4	5	6 (Strongest)
c.	**The warning choices:**	0	1	2	3	4	5	6 (Strongest)
d.	**The destination URL:**	0	1	2	3	4	5	6 (Strongest)
e.	**The website design:**	0	1	2	3	4	5	6 (Strongest)
f.	**Other (please explain):**	0	1	2	3	4	5	6 (Strongest)

9. **Which factor had the most influence on your decision?**

10. **Do you use a computer daily for work? (Circle one)**

Yes No

11. **How many working computers are in your home?**

12. **Rate yourself on this scale regarding computer help:**

I often ask for help 1 2 3 4 5 **Others often ask me for help**

13. **Please list any programming languages that you know:**

14. **Do you have a degree in an IT-related field (e.g. computer science, electrical engineering, etc.)?**

Yes No

15. **Have you attended a computer security conference in the past year?**

Yes No

16. **Have you ever taken or taught a course on computer security?**

Yes No

17. **Is computer security one of your primary job responsibilities?**

Yes No

18. **Please explain what is meant by "phishing":**

19. **Have you ever had any online account information stolen?**

Yes No

20. **Have you ever found fraudulent transactions on a bank statement?**

Yes No

21. **Have you ever had your social security number stolen?**

Yes No

22. **Have you ever been notified that your personal information has been stolen or compromised?**

Yes No

23. **Which web browser do you normally use? (Circle one)**
    a. Internet Explorer
    b. Firefox
    c. Netscape
    d. Safari
    e. Opera
    f. Other: _____

24. **What is your age:_____**
25. **What is your gender? (Circle one)**
    Male Female

26. **What is your highest level of education? (Circle one)**
    a. No high school
    b. Some high school
    c. High school
    d. Some college
    e. College
    f. Professional degree

g. Graduate degree

Appendix E

SSL Warning Study Online Survey

Web Browser Survey

Page One
Thank you for agreeing to further our research. This survey should take less than ten minutes. Please try to be as honest as possible, since there are no right or wrong answers. As our thanks for participating, we will enter you in a drawing for a $75 Amazon.com gift certificate.

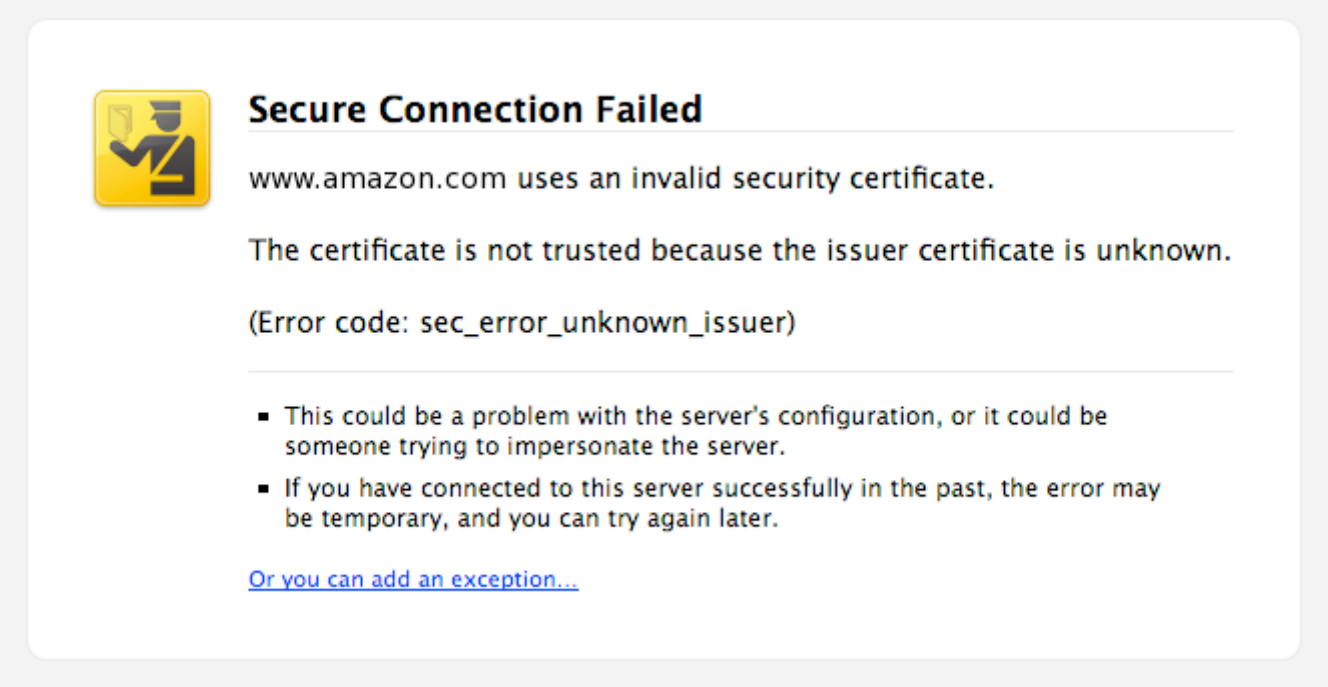
You need to know your web browser version for the next question. You can find it by clicking **here**.
1. Which web browser are you currently using (click **here** to find out)?*

- ( ) Firefox 2
- (•) Firefox 3
- ( ) Internet Explorer 6
- ( ) Internet Explorer 7
- ( ) Safari
- ( ) Opera
- ( ) Other

Click to Next Page
33%

Web Browser Survey

Firefox 3 (unverified)
The questions on this page are about the message below.



19. If you saw this message, would you attempt to continue to the website? *

- ◯ Yes
- ◯ No
- ◯ Other (Please explain)

20. What do you believe this message means? *

21. Have you seen this particular message before? *

- ◯ Yes
- ◯ No
- ◯ I'm not sure

22. If you have seen this message before, please describe the last website that you saw it on:

23. The most recent time that you saw this message, did you continue to the website?*

- Yes
- No
- I haven't seen this message before
- Other (Please explain)

24. How likely is it that something bad would happen if you continued on to the website after seeing this message?*

0% 50% 100%

25. If something bad did happen from continuing on to the website, how bad do you think it would be?*

None Moderate Severe

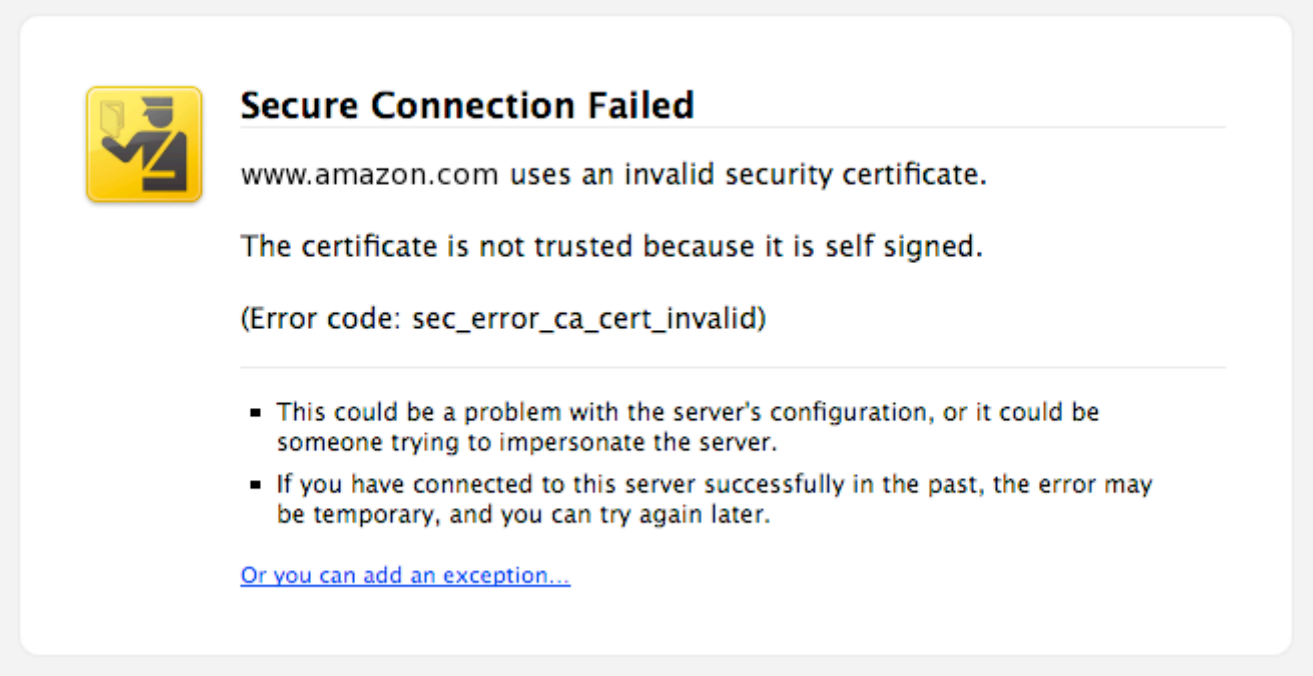
26. What do you believe are the possible consequences of ignoring this message?*

Click to Go Back Click to Next Page

66%

Web Browser Survey

Firefox 3 (self signed)
The questions on this page are about the message below.



51. If you saw this message, would you attempt to continue to the website? *

- ◯ Yes
- ◯ No
- ◯ Other (Please explain)

52. What do you believe this message means? *

53. Have you seen this particular message before? *

- ◯ Yes
- ◯ No
- ◯ I'm not sure

54. If you have seen this message before, please describe the last website that you saw it on:

55. The most recent time that you saw this message, did you continue to the website?*

- Yes
- No
- I haven't seen this message before
- Other (Please explain)

56. How likely is it that something bad would happen if you continued on to the website after seeing this message?*

0% 50% 100%

57. If something bad did happen from continuing on to the website, how bad do you think it would be?*

None Moderate Severe

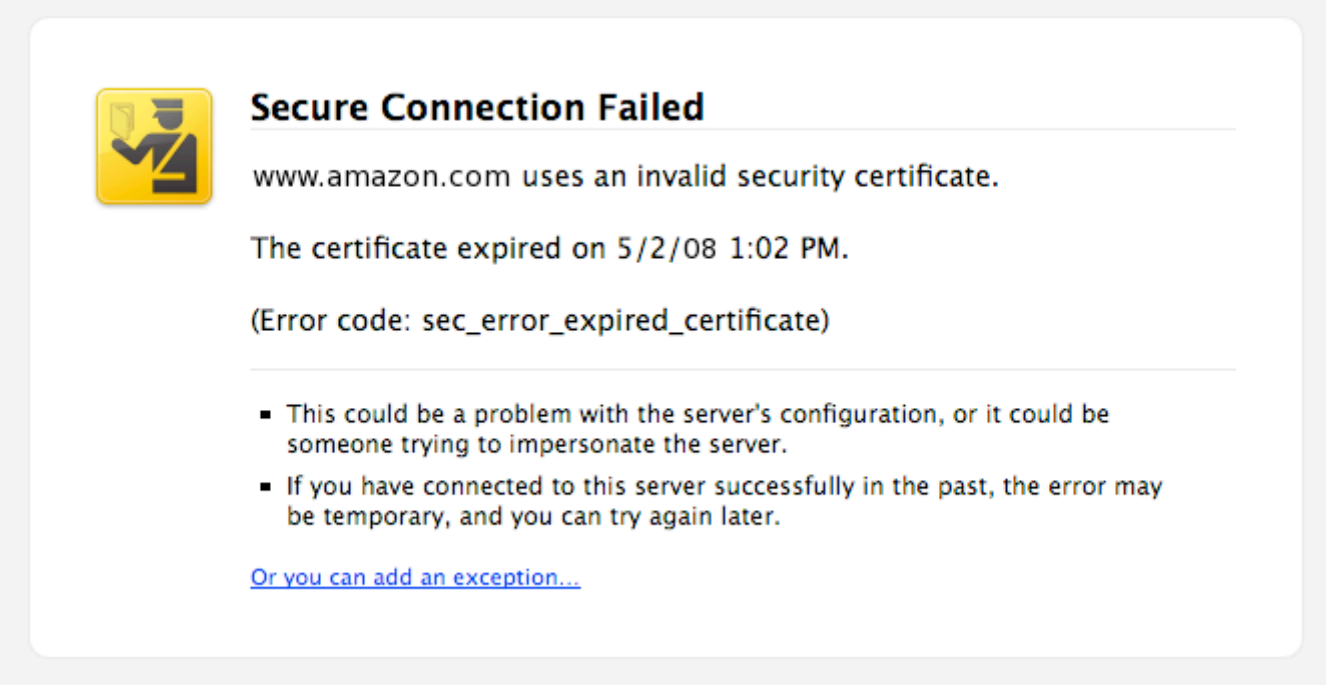
58. What do you believe are the possible consequences of ignoring this message?*

Click to Go Back Click to Next Page

75%

Web Browser Survey

Firefox 3 (expired)
The questions on this page are about the message below.



59. If you saw this message, would you attempt to continue to the website? *

- ◯ Yes
- ◯ No
- ◯ Other (Please explain)

60. What do you believe this message means? *

61. Have you seen this particular message before? *

- ◯ Yes
- ◯ No
- ◯ I'm not sure

62. If you have seen this message before, please describe the last website that you saw it on:

63. The most recent time that you saw this message, did you continue to the website?*

- Yes
- No
- I haven't seen this message before
- Other (Please explain)

64. How likely is it that something bad would happen if you continued on to the website after seeing this message?*

0% 50% 100%

65. If something bad did happen from continuing on to the website, how bad do you think it would be?*

None Moderate Severe

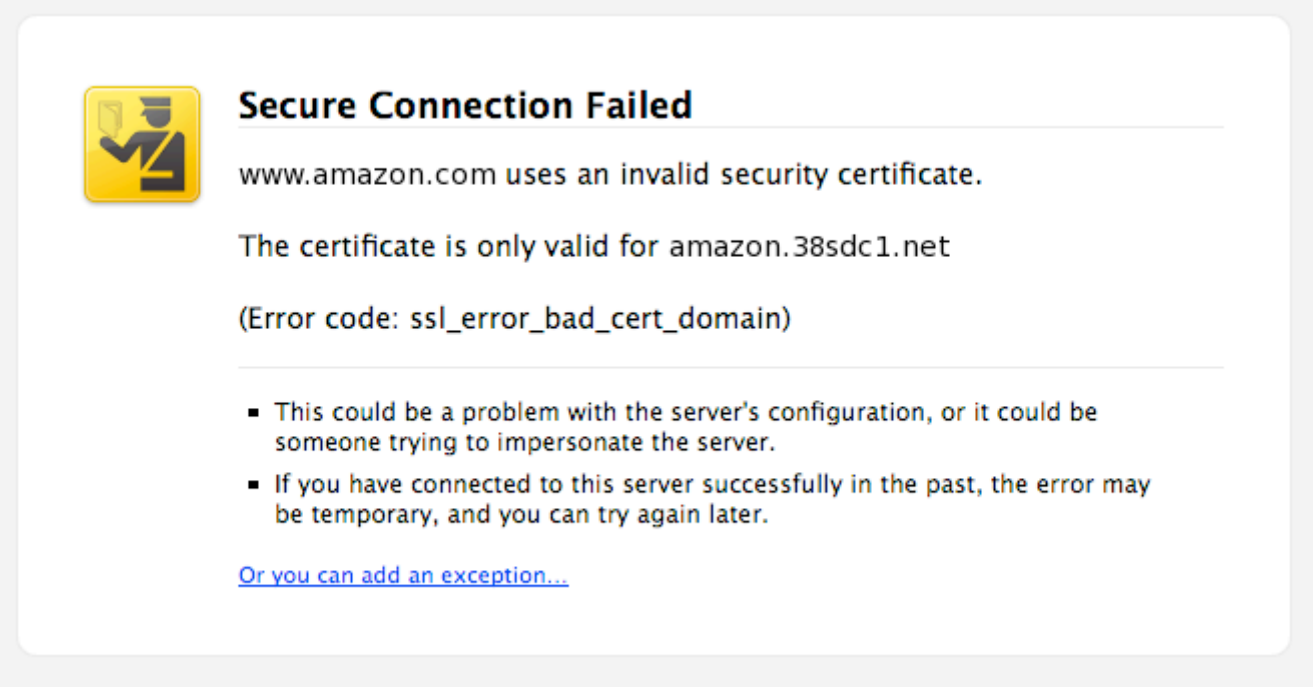
66. What do you believe are the possible consequences of ignoring this message?*

Click to Go Back Click to Next Page

80%

Web Browser Survey

Firefox 3 (mismatch)
The questions on this page are about the message below.



67. If you saw this message, would you attempt to continue to the website? *

- ◯ Yes
- ◯ No
- ◯ Other (Please explain)

68. What do you believe this message means? *

69. Have you seen this particular message before? *

- ◯ Yes
- ◯ No
- ◯ I'm not sure

70. If you have seen this message before, please describe the last website that you saw it on:

71. The most recent time that you saw this message, did you continue to the website?*

- Yes
- No
- I haven't seen this message before
- Other (Please explain)

72. How likely is it that something bad would happen if you continued on to the website after seeing this message?*

0% 50% 100%

73. If something bad did happen from continuing on to the website, how bad do you think it would be?*

None Moderate Severe

74. What do you believe are the possible consequences of ignoring this message?*

Click to Go Back Click to Next Page

83%

Web Browser Survey

Technical Background

In this section we ask you questions about your technical background. Please answer as truthfully as possible, if you don't know what something means, please just say so!

83. Do you use a computer daily for work?*

- Yes
- No

84. How many working computers are in your home?*

85. Where would you rate yourself on this spectrum:*

I often ask others for help with the computer				Others often ask me for help with the computer

86. Do you know any programming languages?*

- Yes
- No

87. If yes, which ones?

88. Do you have a degree in an IT-related field (e.g. computer science, electrical engineering, etc.)?*

- Yes
- No

89. Have you attended a computer security conference in the past year?*

- Yes
- No

90. Have you ever taken or taught a course on computer security?*

- Yes
- No

91. Is computer security one of your primary job responsibilities?*

- Yes
- No

Click to Go Back | Click to Next Page

85%

Web Browser Survey

Demographic Information
You're almost done! Please answer the following demographic information. If you wish to be included in our prize drawing, optionally enter your email address at the bottom.

92. What is your highest level of education?*

- ◯ Less than High School
- ◯ High School/GED
- ◯ Some College
- ◯ 2-year College Degree (e.g. Associates)
- ◯ 4-year College Degree (e.g. BA/BS)
- ◯ Master's Degree (e.g. MA/MS/MBA)
- ◯ Professional Degree (e.g. MD/JD)
- ◯ Doctoral Degree

93. What is your age?*

94. What is your gender?*

- ◯ Male
- ◯ Female

95. How many individuals live in your household?*

96. Have you ever bought anything from Amazon.com?*

- ◯ Yes
- ◯ No

97. Have you ever used Craigslist.org?*

- ◯ Yes
- ◯ No

98. Optionally enter your email address below to be entered in our prize drawing. Your email will only be used to contact you in the event that you win a prize.

Click to Go Back | Finished? Submit your Survey

100%

Appendix F

SSL Warning Study Recruitment Survey

Website Usability Study - Recruitment Survey

Exit this survey >>

Thank you for your interest! This Carnegie Mellon University research study on website usability will give you $10 for roughly 30 minutes of your time. You will be required to show up at our laboratory on the CMU campus. You will receive the full payment on the day of the study. **We are only seeking CMU students, faculty, and staff (i.e. anyone with an Andrew ID).**

***1. Are you still interested in participating in this study?**

- Yes
- No

1 / 4

Next >>

Website Usability Study - Recruitment Survey

Exit this survey >>

Please answer the following questions:

✱2. Indicate how often you use the following websites:

	Never Used	Use 1-10 Times/Year	Use 1-10 Times/Month	Use Daily
Amazon.com	○	○	○	○
eBay.com	○	○	○	○
PayPal.com	○	○	○	○
PNC.com	○	○	○	○
Google.com	○	○	○	○

✱3. Do you have an eBay.com account?

- ○ Yes
- ○ No

✱4. Do you have a PNC.com online banking account?

- ○ Yes
- ○ No

✱5. Do you have a PayPal.com account?

- ○ Yes
- ○ No

✱6. Do you have an Amazon.com account?

- ○ Yes
- ○ No

✱7. Do you have a Google account?

- ☐ Yes
- ☐ No

8. Have you ever participated in a research study at CMU before?

- Yes
- No

9. If yes, what was the purpose of the study?

<< Prev Next >>

Website Usability Study - Recruitment Survey

Exit this survey >>

The contact information that you provide us will only be used for scheduling an appointment for participation in our study. We will not use this information for any other purpose.

***10. What is your name?**

***11. What is your email address?**

***12. What is your Andrew ID?**

***13. What is your phone number?**

***14. Gender:**

- Female
- Male

***15. What is your age?**

16. What is your occupation?

Appendix G

SSL Warning Study Exit Survey

Usability of Information Sources Study

1. Participant Number

∗1. What is your participant number?

Next >>

Usability of Information Sources Study

2. PNC Bank Warning Message

The following 7 Questions (Question 1 to Question 7) are related to the warning you saw at the PNC bank web site.

∗2. Before this study, had you ever seen the warning you saw at the PNC bank web site?

- Yes
- No
- I'm not sure

∗3. Did you read the full text of the warning at the PNC bank web site? Why/why not?

∗4. When the warning at the PNC bank web site was displayed to you, what was your first reaction?

∗5. What did you believe the warning at the PNC bank web site meant?

∗6. After seeing the warning message at the PNC bank web site, did you believe there may be some risk involved with accessing the website?

∗7. What action(s) did you think the warning at the PNC bank web site wanted you to take?

*8. Please explain why you chose to either heed or ignore the warning at the PNC bank web site.

<< Prev Next >>

Usability of Information Sources Study

3. CAMEO Warning Message

The following 7 Questions (Question 8 to Question 14) are related to the warning you saw at the CMU online library catalog (i.e. CAMEO).

∗9. Before this study, had you ever seen the warning at the CMU library catalog?

- Yes
- No
- I'm not sure

∗10. Did you read the full text of the warning at the CMU online library catalog? Why/why not?

∗11. When the warning at the CMU online library catalog was displayed to you, what was your first reaction?

∗12. After seeing the warning message at the CMU library catalog, did you believe there may be some risk involved with accessing the website?

∗13. What action(s) did you think the warning at the CMU online library catalog wanted you to take?

∗14. Did you believe the warning at the CMU online library catalog?

*15. Please explain why you chose to either heed or ignore the warning at the CMU online library catalog.

<< Prev Next >>

Usability of Information Sources Study

4. Security Decision Factors

∗16. How much did the following factors influence your decision to heed or ignore the warnings?

	No Influence at all: 0	1	2	3	4	5	Strongly Influence: 6
The text of the warning							
The colors of the warning							
The choices that the warning presented							
The destination URL							
The look and feel of the destination website							
Other factors (please describe below)							

17. If there were any other factors, please describe them.

∗18. Which factor had the most influence on your decision?

<< Prev Next >>

Usability of Information Sources Study

5. Technical Experience

✱19. Do you use a computer daily for work?

- Yes
- No

✱20. How many working computers are in your home?

✱21. Rate yourself on this scale:

	I often ask others for help with the computer				Others often ask me for help with the computer
Computer help					

✱22. Do you know any programming languages?

- Yes
- No

If yes, which ones?

✱23. Do you have a degree in an IT-related field (e.g. computer science, electrical engineering, etc.)?

- Yes
- No

✱24. Have you attended a computer security conference in the past year?

- Yes
- No

✱25. Have you ever taken or taught a course on computer security?

- Yes

- No

***26. Is computer security one of your primary job responsibilities?**

- Yes
- No

<< Prev Next >>

Usability of Information Sources Study

6. Online Security Questions

∗27. What is a security certificate?

∗28. What is a self-signed certificate?

∗29. Have you ever had any online account information stolen?

- Yes
- No

∗30. Have you ever found fraudulent transactions on a bank statement?

- Yes
- No

∗31. Have you ever had your social security number stolen?

- Yes
- No

∗32. Have you ever been notified that your personal information has been stolen or compromised?

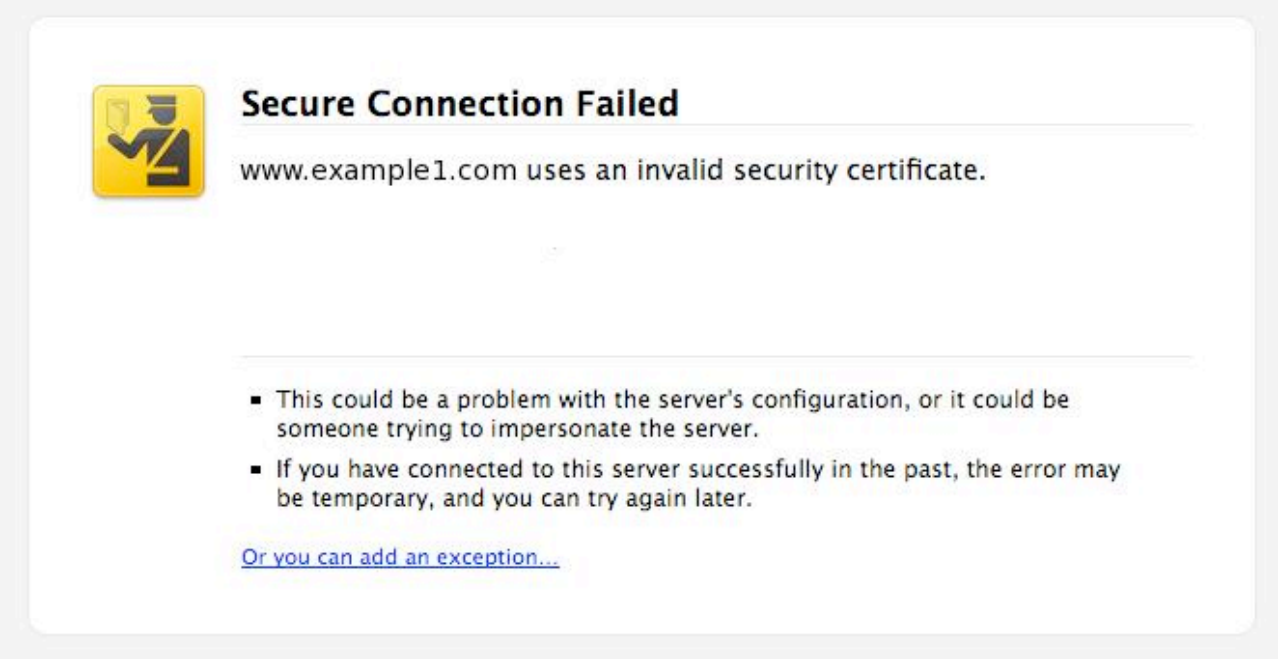
- Yes
- No

<< Prev Next >>

Usability of Information Sources Study

7. Warning Message

The questions on this page are about the message below.



∗33. Imagine you are trying to visit a web site and see the warning message shown above. What does it mean?

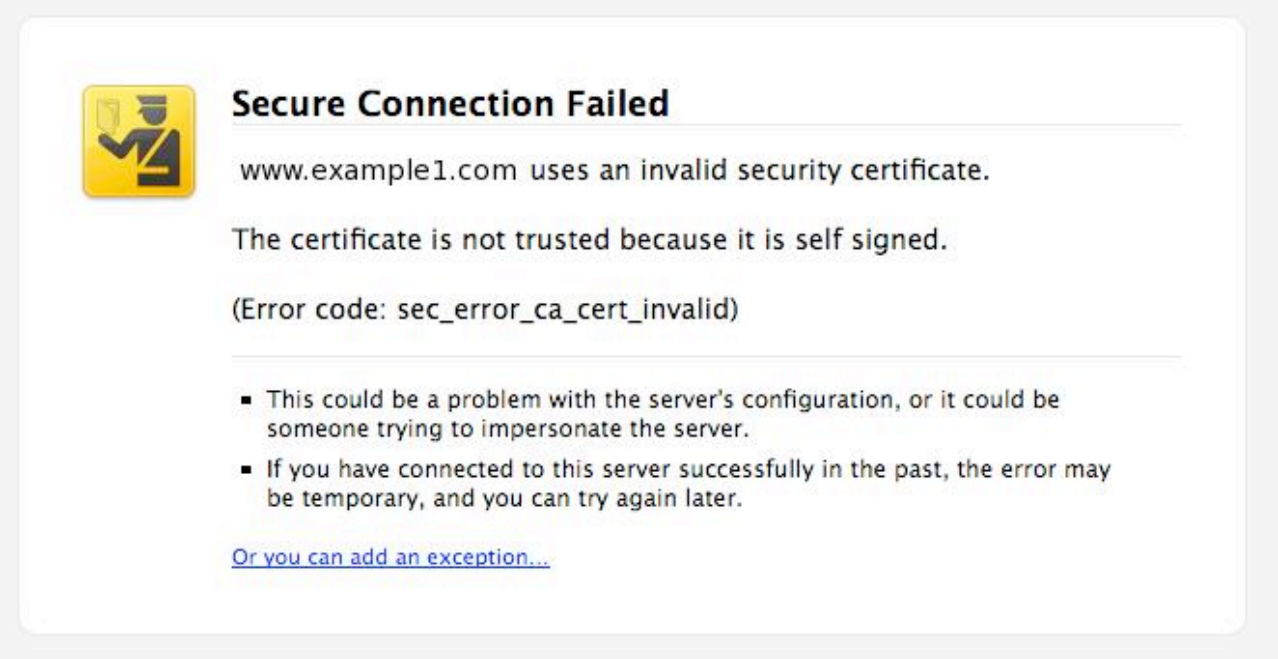
∗34. What would you do if your web browser displayed this message?

<< Prev Next >>

Usability of Information Sources Study

8. Warning Message

The questions on this page are about the message below.



∗35. Imagine you are trying to visit a web site and see the warning message shown above. What does it mean?

∗36. What would you do if your web browser displayed this message?

<< Prev Next >>

Usability of Information Sources Study

9. Demographics

∗37. What Browser do you usually use?

- Internet Explorer
- Firefox
- Netscape
- Safari
- Opera

∗38. What is your age?

∗39. What is your gender?

- Female
- Male

∗40. What is your highest level of education?

- Some high school
- High school diploma
- College degree
- Graduate Degree
- Professional degree (including trade school)
- Other (please specify)

41. What is your position at CMU?

- Ungraduate student
- Graduate student
- Administrative staff
- Faculty

Other (please specify)

*42. What is your department or major?

43. How would you describe your race and ethnicity?

- White
- Black
- Asian or Pacific Islander
- Latino(a)/Hispanic
- Native American
- Other (please specify)

44. What is your country of origin?

45. If you have any additional comments, please write them below.

Thank you for completing this questionnaire! Please raise your hand to notify the study administrator to receive your $10 payment.

<< Prev Done >>

Appendix H

Privacy Information Timing Study Pricing Survey

Online Shopping

Welcome
Thank you for participating in our survey. By completing the survey and providing your e-mail address, you will be entered into the drawing for a **$100 Amazon.com gift certificate**. Your email address will only be used for contacting you in case you win the raffle.

This survey should take approximately 5 minutes to complete. You must be 18 or older to continue.

Click to Next Page

Online Shopping

Introduction
In this survey you will pretend you are using a search engine to purchase an item online. You will be presented with several scenarios of search results. You will be asked to select a website from which you would be most likely to make a purchase using **your own credit card**. Please choose one site, even if this is not a product you would be likely to purchase.

The two products you will be considering are the following:

Batteries (Duracell AA Batteries - 8 pack)

A sex toy (Pocket Rocket Junior - red)

The search results will be presented in a "privacy-enhanced" search engine interface. Websites are rated with "privacy icons" that indicate how good their privacy policies are.

The next page depicts an example of this search engine.

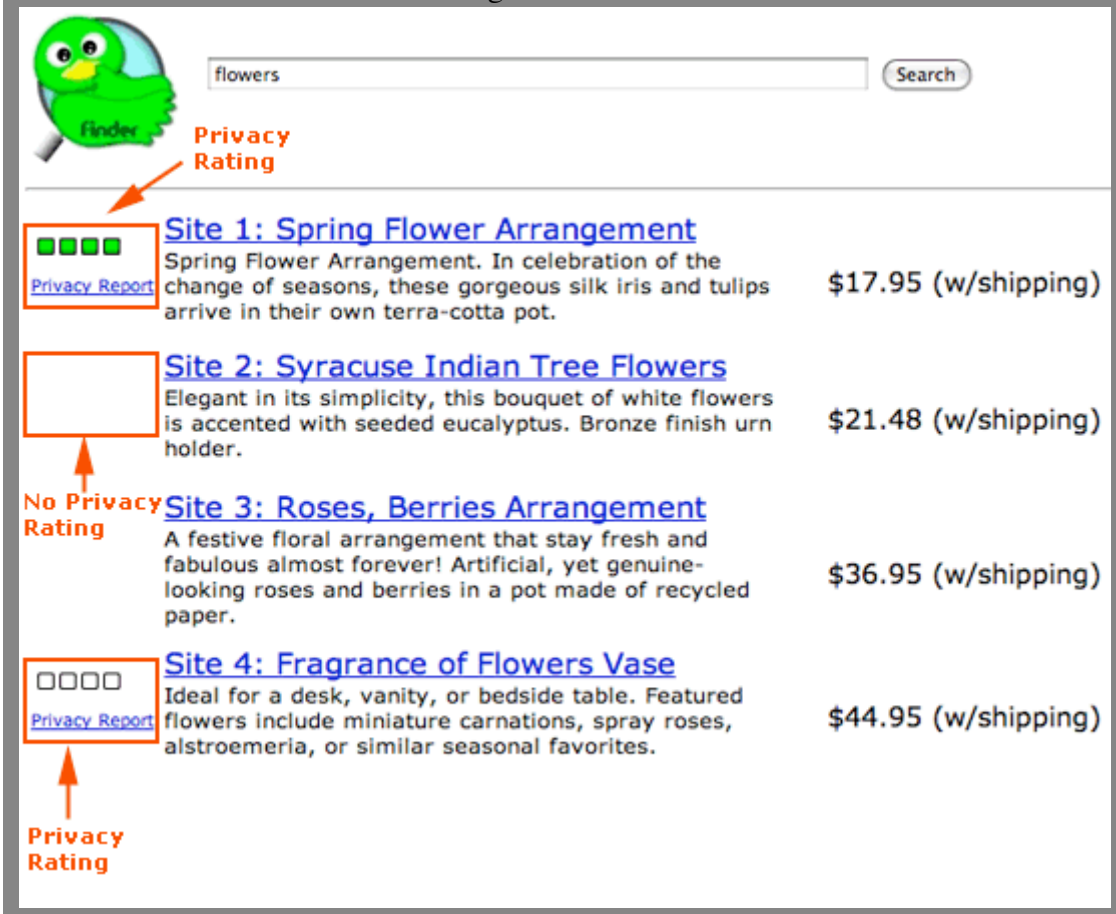
You must answer all questions with a red asterisk.

Click to Go Back Click to Next Page

Online Shopping

Example

This is an example of the search engine results. Please take a moment to familiarize yourself with some of the features of the search engine interface:



Click to Go Back | Click to Next Page

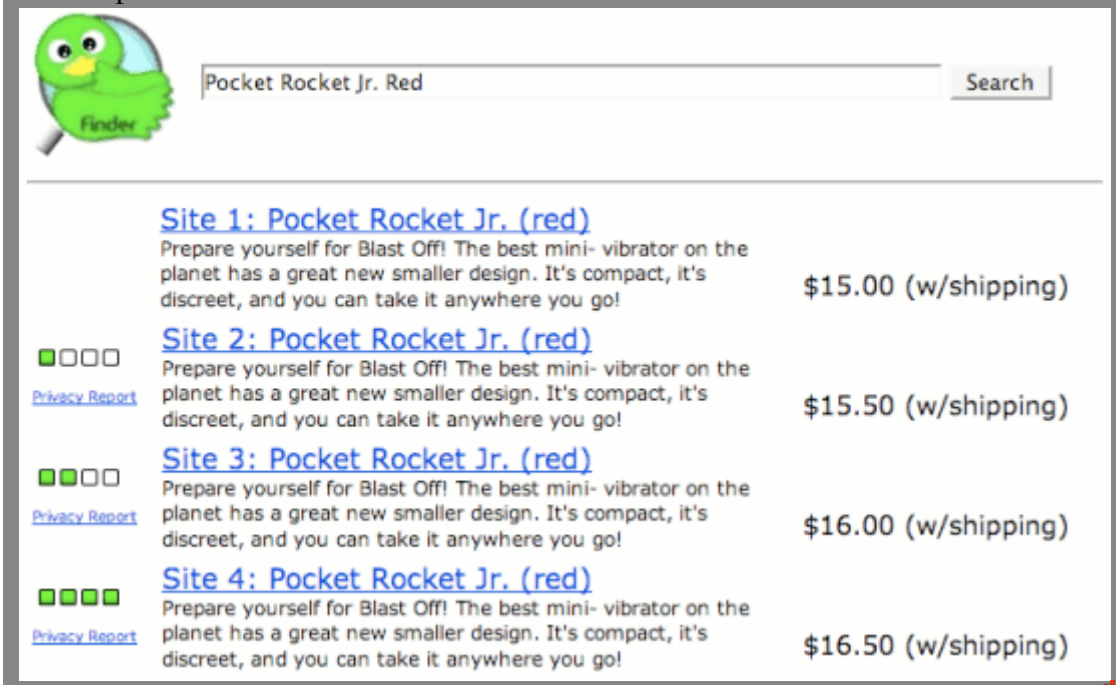
Online Shopping

v7-1

Purchase 1 of 5:

Pretend you are making a purchase for yourself or for a friend using **your own credit card.**
You have just searched for the **Pocket Rocket Jr. Red**.

Given only the information displayed in the search results, from which web site would you be most likely to make this purchase? Please choose one site, even if this is not a product you would be likely to ever purchase.



16. Select the website from which you would be most likely to purchase the **sex toy**.*

- Site 1
- Site 2
- Site 3
- Site 4

Click to Go Back | Click to Next Page

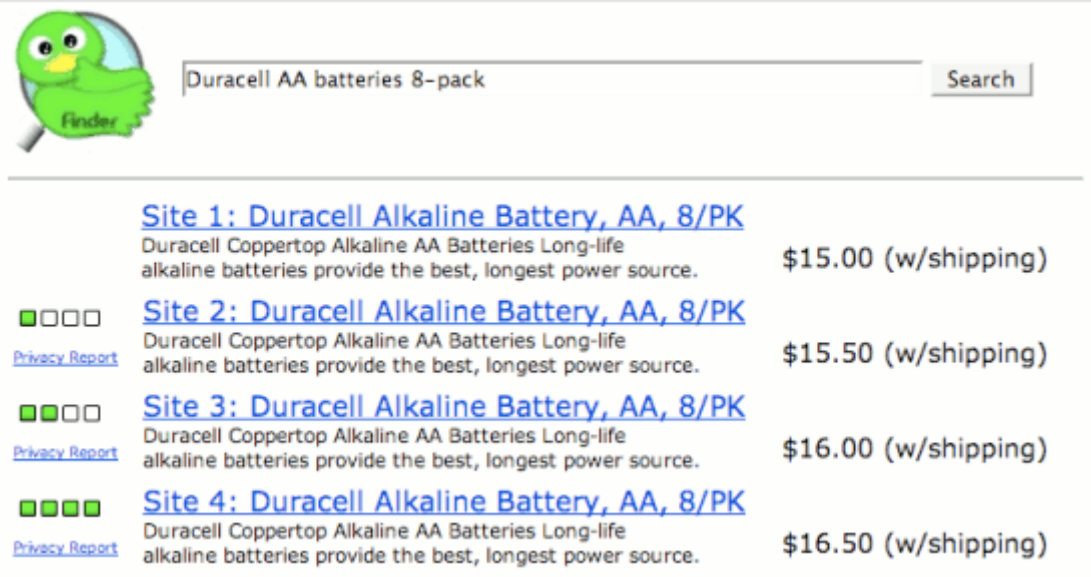
Take a look under the hood

Online Shopping

b7-2
Purchase 2 of 5:

Pretend you are making a purchase for yourself or for a friend using **your own credit card**.
You have just searched for **Duracell AA batteries - 8 pack**.

Given only the information displayed in the search results, from which web site would you be most likely to make this purchase? Please choose one site, even if this is not a product you would be likely to ever purchase.



20. Select the website from which you would be most likely to purchase the **batteries.** *

- ( ) Site 1
- ( ) Site 2
- ( ) Site 3
- ( ) Site 4

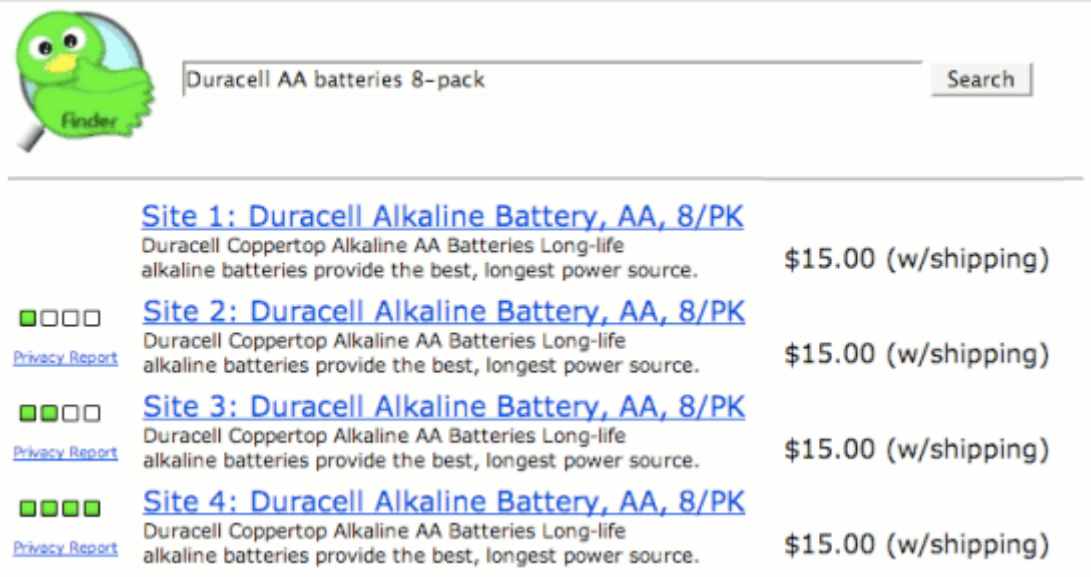
Click to Go Back | Click to Next Page

Online Shopping

bc-3
Purchase 3 of 5:

Pretend you are making a purchase for yourself or for a friend using **your own credit card**.
You have just searched for **Duracell AA batteries - 8 pack**.

Given only the information displayed in the search results, from which web site would you be most likely to make this purchase? Please choose one site, even if this is not a product you would be likely to ever purchase.



23. Select the website from which you would be most likely to purchase the **batteries.**
*

- Site 1
- Site 2
- Site 3
- Site 4

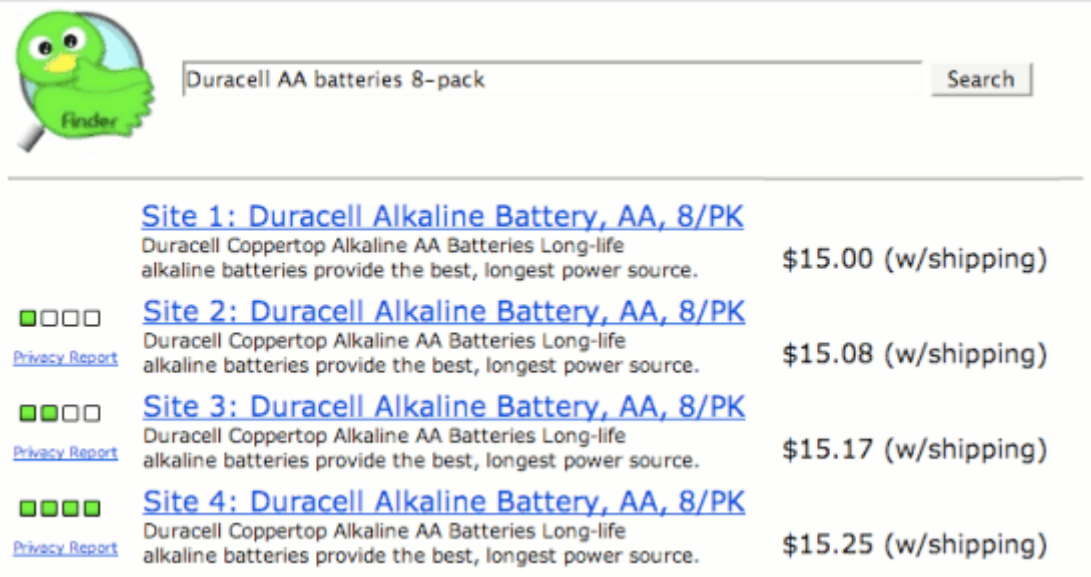
Click to Go Back | Click to Next Page

Online Shopping

b1-4
Purchase 4 of 5:

Pretend you are making a purchase for yourself or for a friend using **your own credit card**.
You have just searched for **Duracell AA batteries - 8 pack**.

Given only the information displayed in the search results, from which web site would you be most likely to make this purchase? Please choose one site, even if this is not a product you would be likely to ever purchase.



25. Select the website from which you would be most likely to purchase the **batteries.**
*

- Site 1
- Site 2
- Site 3
- Site 4

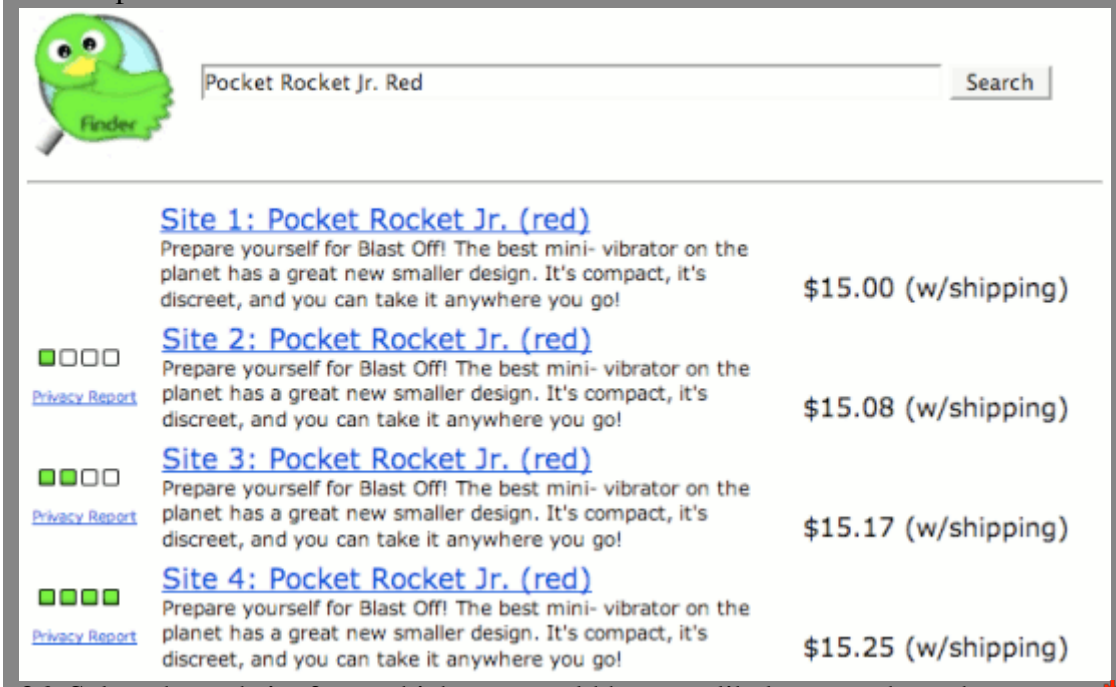
Click to Go Back | Click to Next Page

Online Shopping

v1-5
Purchase 5 of 5:

Pretend you are making a purchase for yourself or for a friend using **your own credit card.**
You have just searched for the **Pocket Rocket Jr. Red**.

Given only the information displayed in the search results, from which web site would you be most likely to make this purchase? Please choose one site, even if this is not a product you would be likely to ever purchase.



26. Select the website from which you would be most likely to purchase the **sex toy**.*

- Site 1
- Site 2
- Site 3
- Site 4

Click to Go Back | Click to Next Page

Take a look under the hood

Online Shopping

Demographic Information
Your survey is almost complete, please enter your **email address** in the box below if you wish to participate in our drawing.

36. What is your gender?*
- Male
- Female

37. What is your age range?*
- Under 18
- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 50
- 51 - 60
- 61 or older

38. What is the highest level of education you've completed?*
- High School
- Vocational Training
- College
- Graduate Program
- Doctorate

39. Have you made a purchase using the Internet in 2008?*
- Yes
- No

Click to Go Back | Finished? Submit your Survey

Appendix I

Privacy Information Timing Study Recruitment Survey

Online Searching and Shopping Study - Recruitment Survey

Exit this survey >>

Thank you for your interest! This Carnegie Mellon University research study on online searching and shopping will give you $50 to shop online for products we specify for you to purchase with your own credit card. You are welcome to keep the change ($15 or more) as well as the products purchased.

You will receive an initial $10 payment on the day of the study and the additional $40 payment after the products you purchased have been shipped.

This study is an "in-person" study, where we will need you to come to a location on the Carnegie Mellon Campus or to Carson St. on the South Side in order to complete the study. We plan on running the study within the next two weeks.

✱1. Are you still interested in participating in this study?

- Yes
- No

Next >>

Online Searching and Shopping Study - Recruitment Survey

Exit this survey >>

***2. What is your name?**

***3. What is your email address?**

***4. What is your phone number?**

***5. Gender:**

- Female
- Male

***6. What is your age?**

7. What is your occupation?

***8. Are you able to come to the CMU campus to participate?**

- Yes
- No

<< Prev Next >>

Online Searching and Shopping Study - Recruitment Survey

Exit this survey >>

You are about halfway done with the survey.

Please answer the following questions:

9. Indicate your level of experience with the following online procedures:

	No Experience	Yearly	Monthly	Weekly	Daily
Using Online Search Engines	○	○	○	○	○
Shopping Online	○	○	○	○	○
Using Instant Messenger Systems	○	○	○	○	○
Checking Email	○	○	○	○	○
Banking Online	○	○	○	○	○

∗10. Have you purchased something online this year?

- ○ Yes
- ○ No

∗11. Do you have an online store/vendor that you often visit or purchase from?

- ○ Yes
- ○ No

12. If yes, what store(s) or vendor(s)?

∗13. When you are selecting a website to purchase an item from, how much do the following factors affect your choice?

	Not at all: 1	2	3	4	5	6	A great deal: 7
Accessibility for sight-impaired users	○	○	○	○	○	○	○

	Not at all: 1	2	3	4	5	6	A great deal: 7
Compatibility of web site with mobile phone web browsers							
Customer Reviews							
Customer Service							
Location of Physical Store							
Page load speed							
Popularity							
Price							
Privacy Policy							
Return Policy							
Software Compatibility							
Shipping Speed							
Website Design							

<< Prev Next >>

Online Searching and Shopping Study - Recruitment Survey

Exit this survey >>

This is the last page of questions for the survey.

✱14. Please rate your level of concern about the following:

	Not concerned at all: 1	2	3	4	5	6	Extremely concerned: 7
A web site uses your health information to determine website content or ads							
A web site shares your health information with other companies							
A website contacts you about other services or products via telephone							
A website contacts you about other services or products via email or postal mail							
A website does not allow you to be removed from marketing/mailing lists							
A website uses your financial information to determine website content or ads							
A website shares your financial information with other companies							

	Not concerned at all: 1	2	3	4	5	6	Extremely concerned: 7
A website does not allow you to find out what information it stores about you							
A website makes its privacy policy available							
A website uses personally identifying information to determine your habits, interests, or other characteristics							
A website shares personally identifying information with other companies							
A website uses information that does not personally identify you to determine your habits, interests, or other characteristics							
A website shares information that does not personally identify you with other companies							

✱15. Please rate the following statements.

	Strongly Disagree: 1	2	3	4	5	6	Strongly Agree: 7
I feel safe giving							

	Strongly Disagree: 1	2	3	4	5	6	Strongly Agree: 7
my personal information to online stores.							
Providing online stores with personal information involves too many unexpected problems.							
I generally trust online companies with handling my personal information and my purchase history.							

***16. Please answer the following question.**

	Not concerned at all: 1	2	3	4	5	6	Extremely concerned: 7
How concerned are you about threats to your personal privacy online in America today?							

You have now completed the survey. You will be contacted shortly to be scheduled for this study.

<< Prev Next >>

Appendix J

Privacy Information Timing Study Exit Survey

CMU Searching and Shopping Study Exit Survey (D)

1. Online Shopping Habits

∗1. How much time do you spend on the Internet per week?

- 1 to 5 hours
- 6 to 10 hours
- 11 to 20 hours
- 21 to 30 hours
- More than 31 hours

∗2. How many online purchases did you make in the last 30 days?

- None
- 1
- 2 or 3
- 4 or 5
- 6 or more

∗3. How much time do you typically spend in an online shopping session when making a purchase?

- Less than 10 minutes
- 11 to 29 minutes
- 30 to 59 minutes
- 1 to 2 hours
- More than 2 hours

Next >>

CMU Searching and Shopping Study Exit Survey (D)

2. Search Engine Specific

∗4. Do you currently use any Internet search engines?

- [ ] Yes
- [ ] No
- [ ] If yes, which one(s)?

∗5. Do you currently use any shopping search engines?

- [ ] Yes
- [ ] No
- [ ] If yes, which one(s)?

∗6. How easy was it to find the following information?

	Very Difficult: 0	1	2	3	4	5	Very Easy: 6
The sizes of reusable bags	○	○	○	○	○	○	○
The color of Ugg boots for women	○	○	○	○	○	○	○
The price of Ugg boots	○	○	○	○	○	○	○
The lifespan of CFLs	○	○	○	○	○	○	○
The CFL replacement for the 100 Watt bulb	○	○	○	○	○	○	○

<< Prev Next >>

CMU Searching and Shopping Study Exit Survey (D)

3. Product Specific Questions

∗7. Before this study, had you ever purchased batteries online?

- Yes
- No

∗8. Was this the first time you made a purchase from the website where you purchased the batteries?

- Yes
- No

∗9. How much did the following factors influence your decision to purchase the batteries from that website?

	No Influence at all: 0	1	2	3	4	5	Strongly Influence: 6
The base price of the product							
The total price of the product (including shipping and taxes)							
The website design or appearance							
Prior experience with the website							
Prior experience with the company (not online)							
The privacy policy							
The return policy							
Other factors (please describe below)							

10. If there were any other factors, please describe them.

∗11. Which factor had the most influence on your decision?

∗12. Before this study, had you ever purchased sex toys online?

- Yes
- No

∗13. Was this the first time you made a purchase from the website where you purchased the Pocket Rocket Jr.?

- Yes
- No

∗14. How much did the following factors influence your decision to purchase the Pocket Rocket Jr. from that website?

	No Influence at all: 0	1	2	3	4	5	Strongly Influence: 6
The base price of the product							
The total price of the product (including shipping and taxes)							
The website design or appearance							
Prior experience with the website							
Prior experience with the company (not online)							
The privacy policy							
The return policy							
Other factors (please describe below)							

15. If there were any other factors, please describe them.

***16. Which factor had the most influence on your decision?**

<< Prev Next >>

CMU Searching and Shopping Study Exit Survey (D)

4. Privacy Preferences

∗17. Please answering the following questions.

	Never: 0	1	2	3	4	5	Always: 6
Do you generally notice whether or not a website you are visiting has a privacy policy?	○	○	○	○	○	○	○
How often do you read websites' privacy policies?	○	○	○	○	○	○	○

∗18. How many privacy policies did you read in the purchasing tasks?

- ○ None of them
- ○ 1
- ○ 2 or 3
- ○ 4 or more

∗19. How much of the privacy policy did you read?
(Check all that apply)

- ☐ None
- ☐ I just clicked the link to make sure there was a privacy policy
- ☐ I skimmed it
- ☐ The first paragraph
- ☐ Half of it
- ☐ The whole thing

20. For the policies you read, why did you read them?

21. For the policies you didn't read, why didn't you read them?

***22. Did you notice the green boxes after you clicked on the URLs?**

- Yes
- No

***23. What did you think the presence of green boxes meant?**

***24. Did the green boxes influence your decision to visit a web site?**

- Yes
- No
- Why or why not?

***25. Did the green boxes influence your decision to purchase from a particular web site?**

- Yes
- No
- Why or why not?

***26. Did you read any of the Privacy Reports provided under the green boxes?**

- Yes
- No

27. If "Yes," what information interested you? (Skip if you answered "No" to the last question.)

- [ ] The location of the website's full privacy policy
- [ ] Conditions under which websites may share your personal information
- [ ] Links to opt-out of additional communications
- [ ] A list of information that is collected about you
- [ ] How your information will be used
- [ ] How you can access your information
- [ ] Company contact information
- [ ] How to resolve privacy-related disputes with the website
- [ ] Other (please specify)

***28. What kinds of things would you expect to find in the privacy policy of a website with four green boxes?**

***29. Would you consider a website with four green boxes to be adequately protecting your privacy?**

- ( ) Yes
- ( ) No
- ( ) I don't know

***30. What kinds of things would you expect to find in the privacy policy of a website with two green boxes?**

***31. Would you consider a website with two green boxes to be adequately protecting your privacy?**

- ( ) Yes

- No
- I don't know

***32. What kinds of things would you expect to find in the privacy policy of a website with boxes where none of the boxes are green?**

***33. Would you consider a website with four empty boxes to be adequately protecting your privacy?**

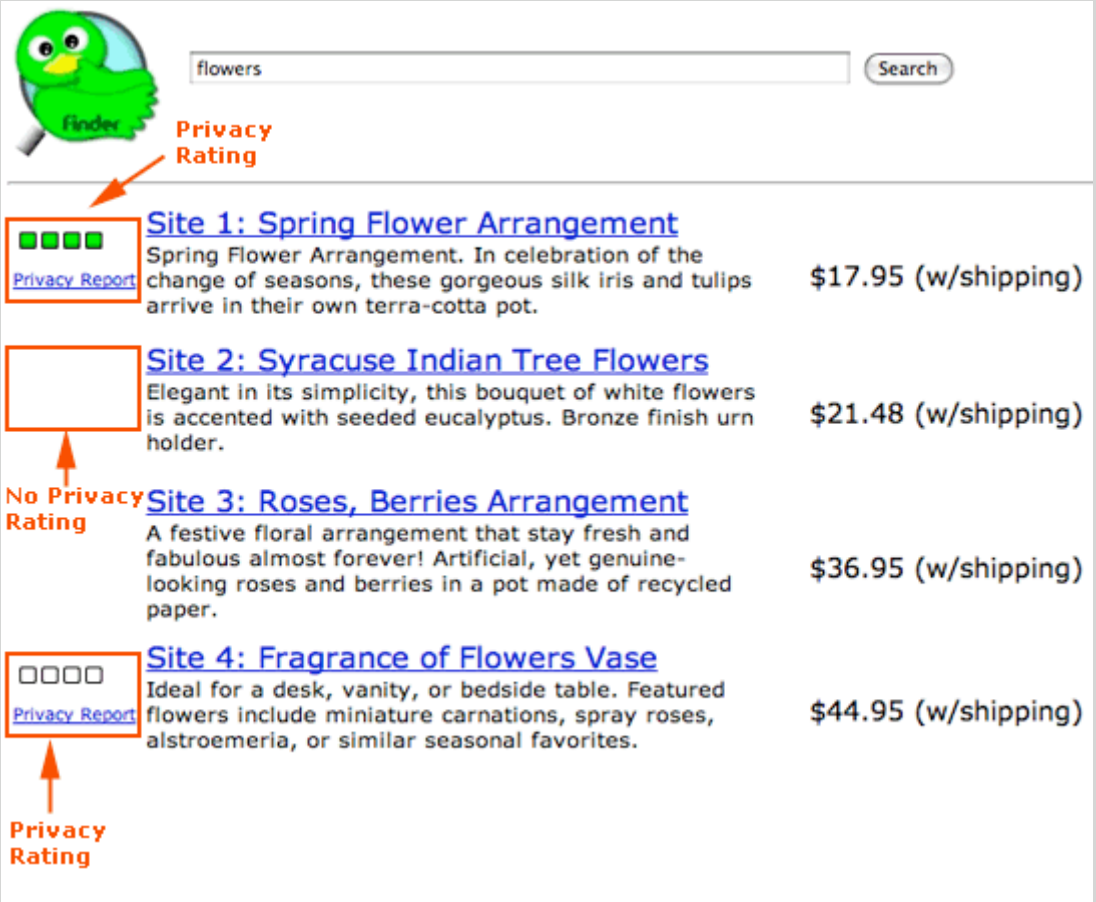
- Yes
- No
- I don't know

<< Prev Next >>

CMU Searching and Shopping Study Exit Survey (D)

5. Privacy Indicators

Please examine the image below.



∗34. What did you think the absence of green boxes means?

∗35. What do you think of the privacy policy of a site with four green boxes compared to a site without any boxes?

The privacy policy of a site with four empty boxes is:

- Better than the one without any boxes
- The same as the one without any boxes
- Worse than the one without any boxes

- I don't know

***36. What do you think of the privacy policy of a site with four empty boxes compared to a site without any boxes?**

The privacy policy of a site with four empty boxes is:

- Better than the one without any boxes
- The same as the one without any boxes
- Worse than the one without any boxes
- I don't know

<< Prev Next >>

CMU Searching and Shopping Study Exit Survey (D)

6. Product Privacy Questions

∗37. What was your level of concern for your privacy when you were purchasing the products in this study?

	Not concerned at all: 0	1	2	3	4	5	Extremely concerned: 6
AA Batteries	○	○	○	○	○	○	○
Sex Toy	○	○	○	○	○	○	○

∗38. When purchasing the batteries, how concerned were you about what the company would do with the following information?

	Not concerned at all: 0	1	2	3	4	5	Extremely concerned: 6
Credit Card Number	○	○	○	○	○	○	○
Email Address	○	○	○	○	○	○	○
Physical Address	○	○	○	○	○	○	○
Phone Number	○	○	○	○	○	○	○
Purchase History	○	○	○	○	○	○	○

∗39. What specific privacy concerns, if any, did you have when purchasing the batteries?

∗40. When purchasing the sex toy, how concerned were you about what the company would do with the following information?

	Not concerned at all: 0	1	2	3	4	5	Extremely concerned: 6
Credit Card Number	○	○	○	○	○	○	○
Email Address	○	○	○	○	○	○	○
Physical Address	○	○	○	○	○	○	○

	Not concerned at all: 0	1	2	3	4	5	Extremely concerned: 6
Phone Number							
Purchase History							

∗41. What specific privacy concerns, if any, did you have when purchasing the sex toy?

∗42. Did you have more concerns with purchasing one product or the other product?

- More concerns when purchasing the AA batteries
- More concerns when purchasing the Pocket Rocket Jr.
- Equal privacy concerns when purchasing both products
- No privacy concerns purchasing either product

∗43. To what extent do you agree or disagree with the following statements?

	Strongly Disagree:0	1	2	3	4	5	Strongly Agree: 6
Consumers have lost all control over how personal information is collected and used by companies							
Most businesses handle the personal information they collect about consumers in a proper and confidential way							
Existing laws and organizational practices provide a reasonable level of protection for							

	Strongly Disagree:0	1	2	3	4	5	Strongly Agree: 6
consumer privacy today							

∗44. Have you ever found fraudulent transactions on your account statement?

- Yes
- No

∗45. Have you ever had your social security number stolen?

- Yes
- No

∗46. Have you ever been notified that your personal information has been stolen or compromised?

- Yes
- No

<< Prev Next >>

CMU Searching and Shopping Study Exit Survey (D)

7. Demographics

∗47. What is your age?

∗48. What is your gender?

- Female
- Male

∗49. What is your highest level of education?

- Some high school
- High school diploma
- College degree
- Graduate Degree
- Professional degree (including trade school)
- Other (please specify)

50. How would you describe your race and ethnicity?

- White
- Black
- Asian or Pacific Islander
- Latino(a)/Hispanic
- Native American
- Other (please specify)

51. What is your country of origin?

Thank you for completing this questionnaire! Please raise your hand to notify the study

administrator and receive your $10 payment. We will send the remainder of your payment ($40) to you after we confirm that the products you ordered have shipped. The study administrator will provide you with instructions for notifying us that your orders have shipped.

<< Prev Done >>

Appendix K

Privacy Finder Usage Study Recruitment Survey

Privacy Finder Usage Study - Recruitment Survey

1 / 2

In order to register for the study, please answer the following questions regarding your privacy preferences.

✱1. Please answer the following questions:

	Never: 0	1	2	3	4	5	Always: 6
Do you generally notice whether or not a website you are visiting has a privacy policy?							
Generally, how often do you read websites' privacy policies?							

✱2. Please rate the following statements.

	Strongly Disagree: 0	1	2	3	4	5	Strongly Agree: 6
I feel safe giving my personal information to online stores.							
Providing online stores with personal information involves too many unexpected problems.							
I generally trust online companies with handling my personal information and							

	Strongly Disagree: 0	1	2	3	4	5	Strongly Agree: 6
my purchase history.							

✱3. Please answer the following question.

	Not concerned at all: 0	1	2	3	4	5	Extremely concerned: 6
How concerned are you about threats to your personal privacy online in America today?							

Next >>

Privacy Finder Usage Study - Recruitment Survey

2 / 2 [redacted]

***4. Gender:**

- Female
- Male

***5. What is your age?**

***6. What is your highest level of education:**

- Some high school
- High school diploma
- Some college
- College degree
- Some professional school
- Professional degree
- Some graduate school
- Graduate degree

***7. What is your country of origin?**

<< Prev Done >>

Bibliography

[1] A. K. Ghosh and C. Howell and J. A. Whittaker. Building software securely from the ground up. IEEE Software, 19(1):14–16, January 2002. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=976936.

[2] M. S. Ackerman, L. F. Cranor, and J. Reagle. Privacy in e-commerce: examining user scenarios and privacy preferences. In EC '99: Proceedings of the 1st ACM Conference on Electronic Commerce, pages 1–8, New York, NY, USA, 1999. ACM.

[3] A. Acquisti. Privacy in electronic commerce and the economics of immediate gratification. In Proceedings of the ACM Electronic Commerce Conference (EC '04), pages 21–29, New York, NY, 2004. ACM Press. http://www.heinz.cmu.edu/ acquisti/papers/privacy-gratification.pdf.

[4] A. Acquisti and J. Grossklags. Privacy and rationality in individual decision making. IEEE Security & Privacy, pages 24–30, January/February 2005. http://www.dtc.umn.edu/weis2004/acquisti.pdf.

[5] A. Adams, M. A. Sasse, and P. Lunt. Making passwords secure and usable. In H. Thimbleby, B. O'Conaill, and P. Thomas, editors, Proceedings of HCI '97, volume People & Computers XII, pages 1–19, Bristol, UK, August 12-15 1997. Springer.

[6] A. Adelsbach, S. Gajek, and J. Schwenk. Visual spoofing of SSL protected web sites and effective countermeasures. In Proceedings of the First Information Security Practice and Experience Conference (ISPEC), pages 204–216, 2005. http://springerlink.metapress.com/openNote.asp?genre=article&issn=0302-9743&volume=3439&spage=204.

[7] W. Adkinson, J. Eisenbach, and T. Lenard. Privacy online: A report on the information practices and policies of commercial web sites. Technical report, Progress & Freedom Foundation, 2002.

[8] E. Agichtein and Z. Zheng. Identifying "best bet" web search results by mining past user behavior. In Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining, (KDD), 2006.

[9] T. S. Amer and J. B. Maris. Signal words and signal icons in application control and information technology exception messages – hazard matching and habituation effects. Technical Report Working Paper Series–06-05, Northern Arizona University, Flagstaff, AZ, October 2006.

[10] Anti-Phishing Working Group. Phishing Activity Trends Report, January, 2007. http://www.antiphishing.org/reports/apwg_report_january_2007.pdf.

[11] A. Anton, J. Earp, Q. He, W. Stufflebeam, D. Bolchini, and C. Jensen. Financial privacy policies and the need for standardization. IEEE Security & Privacy, 2(2):36–45, Mar-Apr 2004.

[12] Bank of America. How Bank of America SiteKey Works for Online Banking Security. http://www.bankofamerica.com/privacy/sitekey/, 2007.

[13] BBC News. Passwords revealed by sweet deal, April 20, 2004. http://news.bbc.co.uk/1/hi/technology/3639679.stm.

[14] M. Beltzner. Bug 399275 – create preference which restores per-page SSL error override option for it professionals. Bugzilla report, Accessed: January 27, 2009. https://bugzilla.mozilla.org/show_bug.cgi?id=399275.

[15] Biju. Bug 398915–secure connection failed at https://mozilla.org/ because of hostname mismatch. Bugzilla report, Accessed: January 27, 2009. https://bugzilla.mozilla.org/show_bug.cgi?id=398915.

[16] J. Birget, D. Hong, and N. Memon. Graphical passwords based on robust discretization. IEEE Transactions on Information Forensics and Security, 1(3):395–399, 2006.

[17] B. Blakely and C. Health. Security Design Patterns. The Open Group, 2004.

[18] S. Brostoff and M. A. Sasse. Are passfaces more usable than passwords? A field trial investigation. In People and Computers XIV–Usability or Else! Proceedings of HCI 2000, Sunderland University, 2000.

[19] J. C. Brustoloni and R. Villamarín-Salomón. Improving security decisions with polymorphic and audited dialogs. In SOUPS '07: Proceedings of the 3rd Symposium on Usable Privacy and Security, pages 76–85, New York, NY, USA, 2007. ACM Press.

[20] S. Byers, L. F. Cranor, and D. Kormann. Automated Analysis of P3P-Enabled Web Sites. In Proceedings of the Fifth International Conference on Electronic Commerce (ICEC2003), pages 197–207, October 1-3, 2003. http://lorrie.cranor.org/pubs/icec03.html.

[21] J. M. Carroll and C. Carrithers. Training wheels in a user interface. Communications of The ACM, 27(8):800–806, 1984.

[22] CBS News. Poll: Privacy Rights Under Attack, October 2, 2005; Accessed: December 17, 2007.

[23] Certification Authority/Browser Forum. Extended validation SSL certificates, Accessed: July 27, 2007. http://cabforum.org/.

[24] C. Chambers, B. Harrison, and J. Vlissides. A debate on language and tool support for design patterns. In POPL '00: Proceedings of the 27th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, pages 277–289, New York, NY, USA, 2000. ACM.

[25] S. Chiasson, R. Biddle, and P. C. van Oorschot. A second look at the usability of click-based graphical passwords. In Proceedings of the 2007 Symposium On Usable Privacy and Security (SOUPS), Pittsburgh, PA, USA, July 18-20 2007.

[26] Install the server certificate. Colorado State online help article, Accessed: January 27, 2009. http://www.engr.colostate.edu/webmail/.

[27] comScore. Inc. comScore Releases January 2009 U.S. Search Engine Rankings, February 18 2009. http://www.comscore.com/press/release.asp?press=2729.

[28] Consumer Reports National Research Center. Consumer Reports Poll: Americans extremely concerned about Internet Privacy, September 25 2008. http://www.consumersunion.org/pub/core_telecom_and_utilities/006189.html.

[29] P. Corporation. Passfaces. http://www.realuser.com/, 2007.

[30] L. F. Cranor. Web Privacy with P3P. O'Reilly and Associates, Sebastopol, CA, 2002.

[31] L. F. Cranor. What do they "indicate?": Evaluating security and privacy indicators. Interactions, 13(3):45–47, 2006. http://doi.acm.org/10.1145/1125864.1125890.

[32] L. F. Cranor. A framework for reasoning about the human in the loop. In Proceedings of the 1st Conference on Usability, Psychology, and Security, Berkeley, CA, 2008. USENIX Association.

[33] L. F. Cranor, S. Byers, D. Kormann, and P. McDaniel. Searching for Privacy: Design and Implementation of a P3P-Enabled Search Engine. In Proceedings of the 2004 Workshop on Privacy Enhancing Technologies (PET2004), pages 314–328, May 26-26, 2004.

[34] L. F. Cranor, P. Guduru, and M. Arjula. User Interfaces for Privacy Agents. ACM Transactions on Computer-Human Interaction, 13(2):135–178, June, 2006.

[35] M. J. Culnan and G. R. Milne. The Culnan-Milne Survey on Consumers and Online Privacy Notices, 2001; Accessed: December 17, 2007. http://intra.som.umass.edu/georgemilne/pdf_files/culnan-milne.pdf.

[36] D. Davis, F. Monrose, and M. Reiter. On user choice in graphical password schemes. In Proceedings of the 13th USENIX Security Symposium, San Diego, CA, 2004.

[37] D. DeJoy, K. Cameron, and L. Della. Postexposure evalution of warning effectiveness: A review of field studies and population-based research. In M. S. Wogalter, editor, Handbook of Warnings, pages 35–48. Lawrence Erlbaum Associates, 2006.

[38] R. E. Dewar. Design and evaluation of public information symbols. In H. J. G. Zwaga and T. Boersma, editors, Visual Information for Everyday Use: Design and Research Perspectives, pages 111–117. Taylor and Francis, London, 1999.

[39] A. J. DeWitt and J. Kuljis. Aligning usability and security: a usability study of polaris. In SOUPS '06: Proceedings of the Second Symposium on Usable Privacy and Security, pages 1–7, New York, NY, USA, 2006. ACM Press.

[40] R. Dhamija and A. Perrig. Dejà vu: A user study using images for authentication. In Proceedings of the 9th USENIX Security Symposium, August 2000. http://citeseer.ist.psu.edu/dhamija00ej.html.

[41] R. Dhamija and J. D. Tygar. The battle against phishing: Dynamic security skins. In Proceedings of the 2005 Symposium on Usable Privacy and Security, New York, NY, USA, July 6-8 2005. ACM Press.

[42] R. Dhamija, J. D. Tygar, and M. Hearst. Why phishing works. In CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 581–590, New York, NY, USA, 2006. ACM Press.

[43] F. Diao and S. S. Sundar. Orienting response and memory for web advertisements: Exploring effects of pop-up window and animation. Communication Research, 31(5):537–567, October 2004.

[44] J. S. Downs, M. Holbrook, and L. Cranor. Decision Strategies and Susceptibility to Phishing. In Proceedings of the 2006 Symposium on Usable Privacy and Security, Pittsburgh, PA, July 12-14, 2006.

[45] J. Earp, A. Anton, L. Aiman-Smith, and W. Stufflebeam. Examining internet privacy policies within the context of user privacy values. IEEE Transactions on Engineering Management, 52(2):227–237, May 2005.

[46] B. Edelman. Adverse selection in online 'trust' certifications. In Proceedings of the 2006 Workshop on the Economics of Information Security (WEIS'06), Cambridge, UK, 2006.

[47] S. Egelman, L. F. Cranor, and A. Chowdhury. An analysis of p3p-enabled web sites among top-20 search results. In Proceedings of the Eighth International Conference on Electronic Commerce, August 14-16, 2006. http://lorrie.cranor.org/pubs/icec06.html.

[48] S. Egelman, L. F. Cranor, and J. Hong. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In CHI '08: Proceeding of the 26th SIGCHI Conference on Human Factors in Computing Systems, pages 1065–1074, New York, NY, USA, 2008. ACM.

[49] S. Egelman, J. Tsai, L. Cranor, and A. Acquisti. Timing is Everything? The Effects of Timing and Placement of Online Privacy Indicators. In CHI '09: Proceeding of the 27th SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 2009. ACM Press.

[50] F. N. Egger. Affective design of e-commerce user interfaces: How to maximise perceived trustworthiness. In Proceedings of The International Conference on Affective Human Factors Design, pages 317–324. Asean Academic Press, 2001.

[51] Federal Trade Commission. National and State Trends in Fraud & Identity Theft. Identity Theft Data Clearinghouse, February 1, 2005. http://www.consumer.gov/sentinel/pubs/Top10Fraud2004.pdf.

[52] E. W. Felten, D. Balfanz, D. Dean, and D. S. Wallach. Web spoofing: An internet con game. In Proceedings of the 20th National Information Systems Security Conference (Baltimore, MD), October, 1997. http://www.cs.princeton.edu/sip/pub/spoofing.html.

[53] A. Ferreira, C. Rusu, and S. Roncagliolo. Usability and security patterns. International Conference on Advances in Computer-Human Interaction, 0:301–305, 2009.

[54] D. Florencio and C. Herley. A large-scale study of web password habits. In WWW '07: Proceedings of the 16th International Conference on the World Wide Web, pages 657–666, New York, NY, USA, 2007. ACM Press.

[55] B. Fogg, J. Marshall, O. Laraki, A. Osipovich, C. Varma, N. Fang, J. Paul, A. Rangekar, J. Shon, P. Swani, and M. Treinen. What Makes Web Sites Credible? A Report on a Large Quantitative Study. In Proceedings of the ACM Computer-Human Interaction Conference, Seattle, WA, March 31 - April 4, 2001. ACM.

[56] S. Fox, L. Rainie, J. Horrigan, A. Lenhart, T. Spooner, and C. Carter. Trust and privacy online: Why Americans want to rewrite the rules. The Pew Internet & American Life Project, August 20, 2000; Accessed: December 17, 2007. http://www.pewinternet.org/pdfs/PIP_Trust_Privacy_Report.pdf.

[57] R. Franco. Better website identification and extended validation certificates in ie7 and other browsers, Accessed: April 4, 2007. http://blogs.msdn.com/ie/archive/2005/11/21/495507.aspx.

[58] B. Friedman, D. Hurley, D. C. Howe, E. Felten, and H. Nissenbaum. Users' conceptions of web security: a comparative study. In CHI '02: CHI '02 Extended Abstracts on Human Factors in Computing Systems, pages 746–747, New York, NY, USA, 2002. ACM Press.

[59] E. Gamma, R. Helm, R. Johnson, and J. Vissides. Design Patterns: Elements of Reusable Object-Oriented Software. Addison Wesley, Reading, MA, 1995.

[60] N. Gandal. The dynamics of competition in the internet search engine market. International Journal of Industrial Organization, 19(7):1103 – 1117, 2001.

[61] S. L. Garfinkel. Design Principles and Patterns for Computer Systems That Are Simultaneously Secure and Usable. PhD thesis, Massachusetts Institute of Technology, 2005.

[62] S. L. Garfinkel, D. Margrave, J. I. Schiller, E. Nordlander, and R. C. Miller. How to make secure email easier to use. In CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 701–710, New York, NY, USA, 2005. ACM.

[63] Gartner, Inc. Gartner Says Number of Phishing E-Mails Sent to U.S. Adults Nearly Doubles in Just Two Years. http://www.gartner.com/it/page.jsp?id=498245, November 9 2006.

[64] E. S. Geller. The Psychology of Safety Handbook. CRC Press, 2 edition, 2001.

[65] E. S. Geller. Working Safe: How to Help People Actively Care for Health and Safety. CRC Press, 2 edition, 2001.

[66] 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.

[67] P. Gutmann. Why isn't the internet secure yet, dammit. In AusCERT Asia Pacific Information Technology Security Conference 2004, Computer Security: Are we there yet?, May 2004. http://conference.auscert.org.au/conf2004/.

[68] C. Haney, C. Banks, and P. Zimbardo. Interpersonal dynamics in a simulated prison. International Journal of Criminology and Penology, 1, 1973.

[69] J. B. Hardee, R. West, and C. B. Mayhorn. To download or not to download: an examination of computer security decision making. Interactions, 13(3):32–37, 2006.

[70] Firefox 3 and secure connections. Online help article, Accessed: January 27, 2009. http://hasylab.desy.de/infrastructure/experiment_control/links_and_tutorials/ff3_and_ssl/index_eng.html.

[71] E. Hellier, D. B. Wright, J. Edworthy, and S. Newstead. On the stability of the arousal strength of warning signal words. Applied Cognitive Psychology, 14:577–592, 2000.

[72] T. Heyman, K. Yskout, R. Scandariato, and W. Joosen. An analysis of the security patterns landscape. In SESS '07: Proceedings of the Third International Workshop on Software Engineering for Secure Systems, Washington, DC, USA, 2007. IEEE Computer Society.

[73] H. Hochheiser. The platform for privacy preference as a social protocol: An examination within the U.S. policy context. ACM Transactions on Internet Technology (TOIT), 2(4):276–306, 2002.

[74] J. Marchesini and S. W. Smith and M. Zhao. Keyjacking: The surprising insecurity of client-side SSL. Computers and Security, 2004. http://www.cs.dartmouth.edu/ sws/papers/kj04.pdf.

[75] C. Jackson and A. Barth. ForceHTTPS: protecting high-security web sites from network attacks. In WWW '08: Proceeding of the 17th International Conference on the World Wide Web, pages 525–534, New York, NY, USA, 2008. ACM.

[76] C. Jackson, D. R. Simon, D. S. Tan, and A. Barth. An evaluation of extended validation and picture-in-picture phishing attacks. In USEC '07: Proceeding of the 1st International Workshop on Usable Security, pages 281–293, Berlin / Heidelberg, Germany, February 2007. Springer.

[77] C. Jensen, C. Sarkar, C. Jensen, and C. Potts. Tracking Website Data-Collection and Privacy Practices with the iWatch Web Crawler. In Proceedings of the 2007 Symposium On Usable Privacy and Security (SOUPS), pages 29–40, Pittsburgh, PA, 2007. ACM Press.

[78] I. Jermyn, A. Mayer, F. Monrose, M. K. Reiter, and A. D. Rubin. The design and analysis of graphical passwords. In Proceedings of the 8th USENIX Security Symposium. The USENIX Association, August 23-26 1999. http://www.ece.cmu.edu/ reiter/papers/1999/USENIX.pdf.

[79] Ka-Ping Yee. Secure Interaction Design and The Principle of Least Authority. In CHI 2003 Workshop on Human-Computer Interaction and Security Systems, April 6, 2003. http://www.andrewpatrick.ca/CHI2003/HCISEC/hcisec-workshop-yee.pdf.

[80] P. Kumaraguru and L. F. Cranor. Privacy Indexes: A Survey of Westin's Studies. Technical Report CMU-ISRI-5-138, Carnegie Mellon University, December, 2005. http://reports-archive.adm.cs.cmu.edu/anon/isri2005/abstracts/05-138.html.

[81] P. Kumaraguru, Y. Rhee, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge. Protecting people from phishing: the design and evaluation of an embedded training email system. In CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 905–914, New York, NY, USA, 2007. ACM Press.

[82] H. Landsberger. Hawthorne Revisited. Cornell University Press, Ithaca, New York, 1958.

[83] A. Lang. The limited capacity model of mediated message processing. Journal of Communication, 50(1):46–70, 2000.

[84] E. Lawrence. IE8 Security Part III: SmartScreen Filter, July 2008. http://blogs.msdn.com/ie/archive/2008/07/02/ie8-security-part-iii-smartscreen-filter.aspx.

[85] S. Lefranc and D. Naccache. Cut-&-paste attacks with Java. In Proceedings of the 5th International Conference on Information Security and Cryptology (ICISC 2002). Springer Berlin/Heidelberg, 2003. http://www.springerlink.com/content/e9j7ajvbc3vn/.

[86] S. Loftesness. Responding to "phishing" attacks, February 23, 2004. http://www.glenbrook.com/opinions/phishing.htm.

[87] M. Madden. Internet Penetration and Impact. April 26, 2006. http://www.pewinternet.org/pdfs/PIP_Internet_Impact.pdf.

[88] C. Masone. Attribute-based, usefully secure email. Technical Report TR2008-633, Dartmouth College, August 2008. PhD Thesis.

[89] A. McDonald and L. Cranor. The cost of reading privacy policies. In Proceedings of the Technology Policy Research Conference, September 26–28 2008.

[90] S. Milgram. Obedience to Authority: An Experimental View. Harpercollins, 1974.

[91] G. R. Milne and M. J. Culnan. Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices. Journal of Interactive Marketing, 18(3):54–61, Summer 2004.

[92] T. Moore and R. Clayton. An empirical analysis of the current state of phishing attack and defence. In Proceedings of the 2007 Workshop on The Economics of Information Security (WEIS2007), May 2007. http://www.cl.cam.ac.uk/ twm29/weis07-phishing.pdf.

[93] T. Moores. Do consumers understand the role of privacy seals in e-commerce? Communications of the ACM, 48(3):86–91, 2005.

[94] mozdev.org. Trustbar, Accessed: April 4, 2007. http://trustbar.mozdev.org/.

[95] C. Nodder. Users and trust: A microsoft case study. In L. F. Cranor and S. Garfinkel, editors, Security and Usability, pages 589–606. O'Reilly and Associates, 2005.

[96] D. A. Norman. The Design of Everyday Things. Doubleday, New York, NY, USA, 1988.

[97] J. Oberheide. Google safe browsing, November 6 2006. http://jon.oberheide.org/blog/2006/11/13/google-safe-browsing/.

[98] OpenDNS. PhishTank Annual Report. http://www.phishtank.com/, October 2007.

[99] A. A. Ozok and S. H. Holden. Alphanumeric and graphical authentication solutions: A comparative evaluation. In Proceedings of HCI International 2005, Las Vegas, Nevada, 2005.

[100] F. D. Paoli, A. L. D. Santos, and R. A. Kemmerer. Vulnerability of "secure" web browsers. In Proceedings of the 20th National Information Systems Security Conference (Baltimore, MD), pages 476–487, 1997.

[101] B. Parno, C. Kuo, and A. Perrig. Phoolproof phishing prevention. In Proceedings of the 10th International Financial Cryptography and Data Security Conference, February 27–March 2, 2006. http://sparrow.ece.cmu.edu/ parno/phishing/docs/phishing.pdf.

[102] A. Patrick. Commentary on research on new security indicators. Self-published online essay, Accessed: January 15, 2009. http://www.andrewpatrick.ca/essays/commentary-on-research-on-new-security-indicators/.

[103] I. Pollach. What's wrong with online privacy policies? Communications of The ACM, 50(9):103–108, 2007.

[104] R. Rasmussen and G. Aaron. Global phishing survey: Domain name use and trends 1h2008. Anti-Phishing Working Group Advisory, November 2008. http://www.antiphishing.org/reports/APWG_GlobalPhishingSurvey1H2008.pdf.

[105] Refsnes Data. Browser statistics, Accessed: April 4, 2007. http://www.w3schools.com/browsers/browsers_stats.asp.

[106] E. Rescorla. SSL and TLS: Designing and Building Secure Systems. Addison Wesley, Reading, MA, 2001.

[107] D. Risney. IE7 http error update. Self-published online tutorial, Accessed: January 15, 2009. http://deletethis.net/dave/xml-source-view/httperror.html.

[108] S. Romanosky, A. Acquisti, J. Hong, L. F. Cranor, and B. Friedman. Privacy patterns for online interactions. In PLoP '06: Proceedings of the 2006 Conference on Pattern Languages of Programs, pages 1–9, New York, NY, USA, 2006. ACM.

[109] B. Ross. Firefox and the worry free web. In L. F. Cranor and S. Garfinkel, editors, Security and Usability: Designing Secure Systems that People Can Use, pages 577–588. O'Reilly Media, Inc., Sebastopol, CA, USA, August 2005.

[110] S. E. Schechter, R. Dhamija, A. Ozment, and I. Fischer. The emperor's new security indicators. In SP '07: Proceedings of the 2007 IEEE Symposium on Security and Privacy, pages 51–65, Washington, DC, USA, 2007. IEEE Computer Society.

[111] B. Schneier. Inside risks: Semantic network attacks. Communications of the ACM, 43(12):168, December 2000. http://doi.acm.org/10.1145/355112.355131.

[112] S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L. Cranor, J. Hong, and E. Nunge. Anti-phishing phil: The design and evaluation of a game that teaches people not to fall for phish. In Proceedings of the 2007 Symposium On Usable Privacy and Security, Pittsburgh, PA, July 18-20, 2007. ACM Press.

[113] E. Sherman. Privacy policies are great—for phds, September 4, 2008. http://industry.bnet.com/technology/1000391/privacy-policies-are-great-for-phds/.

[114] D. K. Smetters and R. E. Grinter. Moving from the design of usable security technologies to the design of useful secure applications. In NSPW '02: Proceedings of the 2002 Workshop on New Security Paradigms, pages 82–89, New York, NY, USA, 2002. ACM Press.

[115] T. Smith-Jackson and M. Wogalter. Methods and procedures in warning research. In M. S. Wogalter, editor, Handbook of Warnings, pages 23–33. Lawrence Erlbaum Associates, 2006.

[116] J. Sobey, R. Biddle, P. C. van Oorschot, and A. S. Patrick. Exploring user reactions to new browser cues for extended validation certificates. In S. Jajodia and J. López, editors, ESORICS: 13th European Symposium on Research in Computer Security, Málaga, Spain, October 6-8, 2008, pages 411–427, 2008. http://dx.doi.org/10.1007/978-3-540-88313-5_27.

[117] E. H. Spafford. Observing reusable password choices. In Proceedings of the 2nd USENIX Security Symposium, September 1992. http://ftp.cerias.purdue.edu/pub/papers/gene-spafford/spaf-OPUS-observe.pdf.

[118] S. Spiekermann and A. Romanow. Attention & interruption management for systems design. Technical report, Institute of Information Systems, Humboldt University Berlin, 2008.

[119] D. W. Stewart and I. M. Martin. Intended and unintended consequences of warning messages: A review and synthesis of empirical research. Journal of Public Policy & Marketing, 13(1):1–1, 1994.

[120] J. Stoll, C. S. Tashman, W. K. Edwards, and K. Spafford. Sesame: informing user security decisions with system visualization. In CHI '08: Proceeding of the 26th SIGCHI Conference on Human Factors in Computing Systems, pages 1045–1054, New York, NY, USA, 2008. ACM.

[121] D. Sullivan. Where Are They Now? Search Engines We've Known & Loved, March 4 2003. http://searchenginewatch.com/2175241.

[122] J. Sunshine, S. Egelman, H. Almuhimedi, N. Atri, and L. F. Cranor. Crying wolf: An empirical study of ssl warning effectiveness. In Proceedings of the 18th USENIX Security Symposium, 2009.

[123] X. Suo, Y. Zhu, and G. Owen. Graphical passwords: A survey. In 21st Annual Computer Security Applications Conference (ACSAC), Tucson, AZ, USA, December 5-9 2005.

[124] F. Tari, A. A. Ozok, and S. H. Holden. A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In Proceedings of the 2006 Symposium on Usable Privacy and Security, Pittsburgh, PA, July 12-14 2006. http://cups.cs.cmu.edu/soups/2006/proceedings/p56_tari.pdf.

[125] The Honeynet Project and The Honeynet Research Alliance. Know Your Enemy: Trend Analysis, 17 December, 2004. http://project.honeynet.org/papers/trends/life-linux.pdf.

[126] J. Tidwell. Designing Interfaces. O'Reilly Media, Inc., 2005.

[127] TRUSTe. TRUSTe Fact Sheet, 2009. http://www.truste.org/about/fact_sheet.php.

[128] J. Tsai, S. Egelman, L. Cranor, and A. Acquisti. The impact of privacy indicators on search engine browsing patterns. Under review.

[129] 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.

[130] J. Turow. Americans and online privacy: The system is broken, June 2003; Accessed: December 17, 2007. http://www.asc.upenn.edu/usr/jturow/internet-privacy-report/36-page-turow-version-9.pdf.

[131] R. Vamosi. Meet Larry, Firefox's friendly passport officer. CNET News, Accessed: January 15, 2009. http://news.cnet.com/8301-10789_3-9970606-57.html.

[132] D. K. van Duyne, J. A. Landay, and J. I. Hong. The Design of Sites: Patterns for Creating Winning Websites. Prentice Hall, 2 edition, 2006.

[133] D. Wendlandt, D. G. Andersen, and A. Perrig. Perspectives: Improving SSH-style host authentication with multi-path probing. In USENIX '08: Proceedings of the 2008 USENIX Annual Technical Conference, Berkeley, CA, USA, June 2008. USENIX Association.

[134] A. F. Westin and H. L. . Associates. Harris-equifax consumer privacy survey (1996). Technical report, Equifax, Inc., 1996.

[135] T. Whalen and K. M. Inkpen. Gathering Evidence: Use of Visual Security Cues in Web Browsers. In Proceedings of the 2005 Conference on Graphics Interface, pages 137–144, Victoria, British Columbia, 2005.

[136] A. Whitten and J. Tygar. Safe staging for computer security. In Proceedings of the 2003 Workshop on Human-Computer Interaction and Security Systems, April 2003.

[137] 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.

[138] S. Wiedenbeck, J. Waters, J.-C. Birget, A. Brodskiy, and N. Memon. Passpoints: Design and longitudinal evaluation of a graphical password system. International Journal of Human Computer Studies, 63:102–127, 2005.

[139] M. Wogalter. Purpose and scope of warnings. In M. Wogalter, editor, Handbook of Warnings, pages 3–9. Lawrence Erlbaum Associates, Mahway, NJ, USA, 2006.

[140] M. S. Wogalter. Communication-Human Information Processing (C-HIP) Model. In M. S. Wogalter, editor, Handbook of Warnings, pages 51–61. Lawrence Erlbaum Associates, 2006.

[141] M. S. Wogalter, M. Kalsher, L. J. Frederick, A. B. Magurno, and B. M. Brewster. Hazard level perceptions of warning components and configurations. International Journal of Cognitive Ergonomics, 2:123–143, 1998.

[142] M. S. Wogalter and S. D. Leonard. Attention capture and maintenance. In M. S. Wogalter, D. M. DeJoy, and K. R. Laughery, editors, Warnings and Risk Communication, chapter 7. CRC Press, 1999.

[143] M. S. Wogalter and N. C. Silver. Arousal strength of signal words. Forensic Reports, 3:407–420, 1990.

[144] M. S. Wogalter and W. J. Vigilante. Attention switch and maintenance. In M. S. Wogalter, editor, Handbook of Warnings, pages 245–265. Lawrence Erlbaum Associates, New Jersey/London, 2006.

[145] M. Wu. Fighting Phishing at the User Interface. PhD thesis, Massachusetts Institute of Technology, August 2006.

[146] M. Wu, R. C. Miller, and S. L. Garfinkel. Do security toolbars actually prevent phishing attacks? In CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 601–610, New York, NY, USA, 2006. ACM.

[147] H. Xia and J. C. Brustoloni. Hardening web browsers against man-in-the-middle and eavesdropping attacks. In WWW '05: Proceedings of the 14th International Conference on the World Wide Web, pages 489–498, New York, NY, USA, 2005. ACM.

[148] J. Yan, A. Blackwell, R. Anderson, and A. Grant. The memorability and security of passwords–some empirical results. Technical Report UCAM-CL-TR-500, University of Cambridge, 15 JJ Thomson Avenue, Cambridge CB3 0FD, United Kingdom, September 2000. http://www.cl.cam.ac.uk/TechReports/UCAM-CL-TR-500.pdf.

[149] Z. E. Ye and S. Smith. Trusted paths for browsers. In Proceedings of the 11th USENIX Security Symposium, pages 263–279, 2002. http://www.usenix.org/events/sec02/full_papers/ye/ye.pdf.

[150] Z. E. Ye, S. Smith, and D. Anthony. Trusted paths for browsers. ACM Transactions on Information and System Security (TISSEC), 8(2):153–186, 2005.

[151] K.-P. Yee and K. Sitaker. Passpet: Convenient password management and phishing protection. In SOUPS '06: Proceedings of the Second Symposium on Usable Privacy and Security, pages 32–43, New York, NY, USA, 2006. ACM Press. http://doi.acm.org/10.1145/1143120.1143126.

[152] J. Yoder and J. Baraclow. Architectural patterns for enabling application security. In Proceedings of Pattern Languages of Programs (PLoP), 1997.

[153] S. L. Young and M. S. Wogalter. Memory of intruction manual warnings: Effect of pictorial icons and conspicuous print. In Proceedings of the Human Factors Society 32nd Annual Meeting, pages 905–909, 1988.

[154] Y. Zhang, S. Egelman, L. F. Cranor, and J. Hong. Phinding phish: Evaluating anti-phishing tools. In Proceedings of the 14th Annual Network & Distributed System Security Symposium (NDSS 2007), 28th February - 2nd March, 2007. http://lorrie.cranor.org/pubs/toolbars.html.

[155] M. E. Zurko and R. T. Simon. User-centered security. In NSPW '96: Proceedings of the 1996 Workshop on New Security Paradigms, pages 27–33, New York, NY, USA, 1996. ACM Press.

[156] How do I configure ZXTM with Firefox 3? ZXTM KnowledgeHub, Accessed: January 27, 2009. http://knowledgehub.zeus.com/faqs/2008/02/05/configuring_zxtm_with_firefox_3.