# Crying Wolf: An Empirical Study of SSL Warning Effectiveness

Joshua Sunshine, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor
Carnegie Mellon University
{sunshine, egelman, hazim}@cs.cmu.edu, natri@andrew.cmu.edu, lorrie@cs.cmu.edu

## Abstract

Web users are shown an invalid certificate warning when their browser cannot validate the identity of the websites they are visiting. While these warnings often appear in benign situations, they can also signal a man-in-the-middle attack. We conducted a survey of over 400 Internet users to examine their reactions to and understanding of current SSL warnings. We then designed two new warnings using warnings science principles and lessons learned from the survey. We evaluated warnings used in three popular web browsers and our two warnings in a 100-participant, between-subjects laboratory study. Our warnings performed significantly better than existing warnings, but far too many participants exhibited dangerous behavior in all warning conditions. 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.

## 1 Introduction

Browsers display Secure Socket Layer (SSL)[1] warnings to warn users about a variety of certificate problems, for example when the server's certificate has expired, mismatches the address of the server, or is signed by an unrecognized authority. These warning messages sometimes indicate a man-in-the-middle or DNS spoofing attack. However, much more frequently users are actually connecting to a legitimate website with an erroneous or self-signed certificate.

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 [19, 22]. In this paper we examine user reactions to five different SSL 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. Finally, by asking users a question they can answer, the system can tailor a warning to the user's situation and instruct users in the appropriate steps necessary to avoid any hazard.

We conducted a survey of 409 Internet users' reactions to current web browser SSL warnings and found that risk perceptions were the leading factor in respondents' decisions of whether or not to visit a website with an SSL error. However, those who understood the risks also perceived some common SSL warnings as not very risky, and were more likely to override those warnings.

[1] The Secure Socket Layer (SSL) and Transport Layer Security (TLS) protocols secure web communication by encrypting data sent between browser and server and by validating the identity of the server. For the remainder of the paper we will use the common convention of using the term "SSL" to refer to both protocols.

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.

In the next section we provide an overview of other studies that have been conducted on web browser security indicators. In Section 3 we present our online SSL warning survey methodology and results. In Section 4 we present our laboratory experiment methodology and results. Finally, we discuss our findings and conclusions.

## 2 Background and Related Work

Much previous research has indicated that users do not understand SSL. A study in 2002 found that half of the participants could not identify a secure browser connection [8]. A 2005 study tracked eye movements and found that participants paid no attention to web browser security cues such as SSL icons. Only after priming participants to be on the lookout for security information, 69% of participants noticed the lock icon [21]. Schechter et al. tested the usability of security indicators by removing SSL indicators from a banking website and observed that all 63 participants still provided their passwords [17].

The major web browsers now include support for extended validation (EV) certificates. 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 corporate entity [2]. FF3 and IE7 indicate a website has an EV certificate by coloring the address bar green and displaying the name of the website owner. However, a study by Jackson et al. 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, after reading a help file, users were less suspicious of fraudulent websites that did not yield warning indicators [11]. Sobey et al. performed an eye tracking study in 2008 to examine whether participants would notice simulated versions of the EV certificate indicators that are used by FF3. They found that none of their 28 participants examined the address bar when making online shopping decisions, and therefore none of them encountered the secondary SSL dialogs containing information about the website owners [18].

Usability problems with security indicators in web browsers go beyond SSL. Wu et al. conducted a study of security toolbars used to help users identify phishing websites. The researchers examined three different styles of passive indicators—indicators that do not force user interactions—that appeared in the browser chrome. They discovered that 25% of the participants failed to notice the security indicators because they were focused on the primary task. In fact, 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 [23]. The factors that go into website trust have been exten-

sively studied by Fogg et al., who found that the "look and feel" of a website is often most important for gaining user trust [7]. Thus users might trust a professional looking website despite the presence of a passive security indicator. Dhamija et al. corroborated these findings by performing a study on why people fall for phishing websites. In their study, users examined a set of websites and were asked to identify which ones were phishing websites. They found that 23% of their study participants did not look at any of the web browser security indicators when making their decisions, even though the participants were primed for security. The researchers concluded that passive security indicators are ineffective because they often go unnoticed [4].

Because of the problems with passive security indicators, many web browsers now display "active" warnings that require the user to take an action—usually deciding whether or not to visit the destination website—in order to dismiss the warning. While these warnings force the user to acknowledge them, they still allow the user to ignore their advice and proceed to the website despite the security error. In 2008, Egelman et al. performed a study on active web browser warnings used to warn users about potential phishing websites. They discovered that users who claimed to have seen the warnings before were significantly more likely to ignore them in the laboratory. They concluded that many of the participants had become habituated to seeing similar-looking warnings when browsing legitimate websites, and are now likely to ignore all future similarly-designed warnings, regardless of the danger they represent [6].

Jackson and Barth address the problem of users ignoring SSL warnings with the ForceHTTPS system [10]. Websites with CA signed certificates deploy a special ForceHTTPs cookie to a user's browser, which from then on only accepts valid SSL connections to the website. This strategy is elegantly simple, but it does not protect users when they encounter a website for the first time.

Wendlandt et al. created the Perspectives system to prevent habituation by only displaying warnings when an attack is probable. Perspectives transforms the CA model into a "trust-on-first-use" model, similar to how SSH works. "Notaries" keep track of all previously viewed SSL certificates and only warn users when they encounter a certificate that has changed over time. This eliminates many common SSL errors, thereby only displaying warnings when an attack is probable [20]. However, when users do encounter certificates that have been altered, it is unclear how the warnings should be designed so as to maximize their effectiveness.

Xia and Brustoloni implement a system to help users better react to unverified certificates [24]. The system requires websites interested in using private CA signed certificates to distribute tokens to their users by physical media. In 2007, Brustoloni and Villamarín-Salomón explored the idea of creating polymorphic dialogs to combat habituation. While their preliminary results were promising for warning users about malicious email attachments, it is unclear what the long-term efficacy would be if such a system were created for SSL warnings [1].

The pervasive nature of SSL errors raises questions about the efficacy of SSL warnings. A survey of 297,574 SSL-enabled websites queried in January 2007 found 62% of the websites had certificates that would trigger browser warnings [5]. A January 2009 study performed using a list of the top one million websites found that at least 44% of the 382,860 SSL-enabled websites had certificates that would trigger warnings [13].[2] Given this large sample, many of the errors may appear on websites that are not frequently visited. Our own analysis of the top 1,000 SSL-enabled websites yielded 194 SSL errors, which is still an alarming number. Unfortunately, we do not have data on the proportion of certificate errors that appear on legitimate websites versus malicious websites, making it unclear whether these particular errors are indicative of an ongoing attack. However, we believe it is likely that most certificate errors occur on non-malicious websites, and therefore many users view the associated warnings as false positives. This means that if a web browser displays a particular warning each time it encounters any type of certificate error, users will quickly become habituated to this warning regardless of the underlying error.

[2]This estimate is likely low as the 2009 study did not catalog domain name mismatch errors.

# 3 SSL Survey

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.

## 3.1 Methodology

We presented survey respondents with screenshots of three different SSL warnings from the browser that they were using at the time they took the survey[3] and asked them several questions about each warning. These questions were followed by a series of questions to determine demographic information.

We showed participants warnings for expired certificates, certificates with an unknown issuer, and certificates with mismatched domain names.[4] 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 to determine whether they understood what the warnings mean, what they would do when confronted with each warning, and their beliefs about the consequences of ignoring these warnings.

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, 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.

## 3.2 Analysis

Our 409 survey respondents used the following browsers: 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 no significant 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. 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.

[3] We used screenshots of the warnings from FF2, FF3, and IE7. Users of web browsers other than FF2, FF3, or IE7 were only asked the demographic questions.

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

[5] All statistics were evaluated with $\alpha$=0.05. We used a Fisher's exact test for all statistics where we report a p-value only.

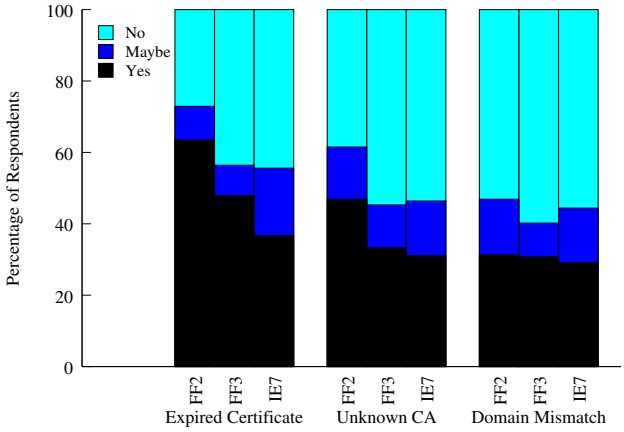



Figure 1: Participant responses to the question: *If you saw this message, would you attempt to continue to the website?*

### 3.2.1 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 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 correct. As shown in Table 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 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.

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 web browser vulnerabilities or trick visitors into downloading malware, we considered these outside the scope of our survey because they either impact only users of a specific browser 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]."

| 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 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.

> "On my Linux box, nothing significantly bad would happen."

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." We found significant differences in responses to each warning for all three web browsers: respondents consistently ranked the expired certificate warning as being less risky than both of the other warnings. Table 2 depicts the perceived likelihood of risk for each of the web browsers and each of the three SSL warnings.

To examine whether there were differences in risk perception based on the underlying SSL error, 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 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.

#### 3.2.2 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 3.1, 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 4 lists the average scores for each of the web browser conditions. We classified those with tech scores greater than or equal to two as "experts." The expert group 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 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-

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

Table 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 | $\chi^2_2 = 20.49$ | $p < 0.0005$ |
| **FF3** | 1.96 | 2.36 | 2.32 | $\chi^2_2 = 9.00$ | $p < 0.011$ |
| **IE7** | 2.14 | 2.36 | 2.34 | $\chi^2_2 = 16.90$ | $p < 0.0005$ |

Table 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.

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.

#### 3.2.3 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. 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.

# 4 Laboratory Experiment

We conducted a laboratory study to determine the effect of SSL warnings on user behavior during real tasks.

## 4.1 Methodology

We designed our laboratory study as a between-subjects experiment with five conditions: FF2 (Figure 2(a)), FF3 (Figure 3), IE7 (Figure 2(b)), a single-page redesigned warning (Figure 4(b)), and a multi-page redesigned warning (Figure 4). We asked participants to find information using four different types of information sources. Each task included a primary information source—a website—and an alternate source that 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.[6] 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

[6]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.

| | 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 4: Percentage of experts and non-experts who said they would continue past the warnings. The first column shows respondents' average tech scores.

the alternative information source. Finally, we gave users an exit survey to gauge their understanding of and reaction to the warnings.

#### 4.1.1 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.[7] All recruits were given an online screening survey, 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. 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 whose responses were not used to screen participants (e.g. "How often do you use Amazon.com?"), to further obfuscate the study purpose.

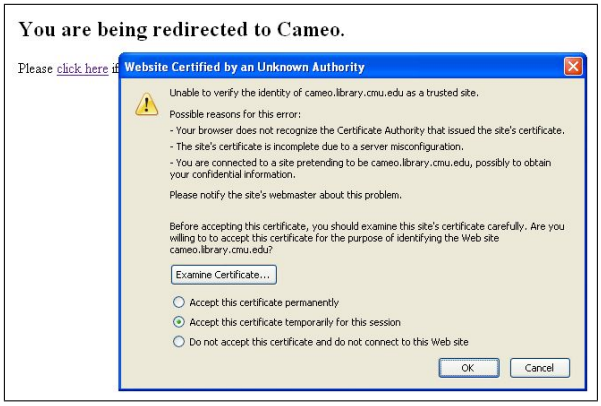
[7]Initially participants were paid \$10, but we raised the payment to \$20 to reach our recruiting goals.

#### 4.1.2 Conditions

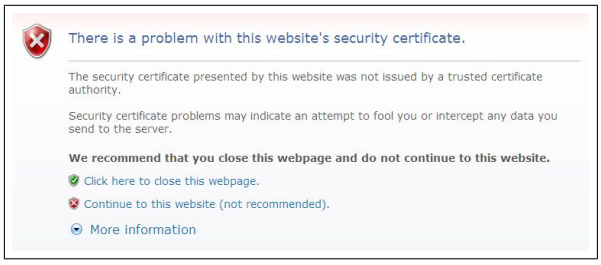
The FF2 warning, displayed in Figure 2(a), is typical of invalid certificate warnings prior to 2006. This warning has a number of design flaws. The text contains jargon such as, "the website's certificate is incomplete due to a server misconfiguration." The look and feel of the warning, a grey dialog 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 the online banking application in our study.

A more subtle problem with the FF2 warning, and those like it, is that it asks users a question that they cannot answer. 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, in the words of Firefox project co-founder Blake Ross, "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 [16].

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




(a) Firefox 2




(b) Internet Explorer 7

Figure 2: Screenshots of the FF2 and IE7 warnings.

the user has to click a link, mysteriously labeled "or you can add an exception..." (Figure 3), then click a button that opens a dialog requiring two more button clicks. The first version of the FF3 warning required 11 steps.[8] 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.[9]

The IE7 warning, shown in Figure 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

[8]https://bugzilla.mozilla.org/show_bug.cgi?id=399275
[9]https://bugzilla.mozilla.org/show_bug.cgi?id=398915

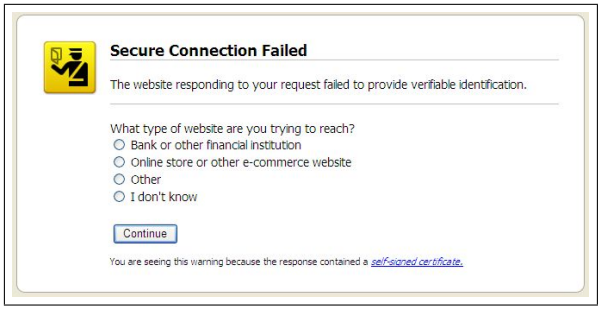

Figure 3: Screenshot of the initial FF3 warning.

can be overridden with a single click on a link labeled "Continue to this website." It has a slightly scarier look 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 4(a), and then, depending on the response, delivers the user either to the severe warning page shown in Figure 4(b) or to the requested website. The second version of the warning shows only the severe warning (Figure 4(b)). Both versions were implemented in IE7. We used the resourcemodify tool[10] 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 the Firefox "passport officer."[11] 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

[10]http://deletethis.net/dave/xml-source-view/httperror.html
[11]http://news.cnet.com/8301-10789_3-9970606-57.html




(a) Page 1


(b) Page 2

Figure 4: Screenshot of redesigned warning.

default button, 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 questions, and instead decided 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 websites use invalid certificates. Instead, our warning aims to ask the users a question that they can answer and will allow us to assess the risk level. 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 website," "other," and "I don't know." If users selected the first two options, they saw the severe warning that discouraged them from continuing. We tested this question as a prototype for leveraging user-provided information to improve security warnings. It is not a complete solution as our question neglects many other types of websites that may collect sensitive information. We decided to show the secondary warning on bank websites and online stores because these are the most frequently attacked websites [15].

#### 4.1.3 Experimental Setup

All studies were conducted in our laboratory on the same model of laptop. Participants interacted with the laptop within a virtual machine (VM). We reset the VM to a snapshot after each participant finished the study to destroy any sensitive data entered by the participant (e.g. bank password). This process also ensured that all browser and operating system settings were exactly the same for every participant. Finally, experimenters read instructions to participants from a script and experimenters did not help particiants complete the tasks.

#### 4.1.4 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 next task was not 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* (i.e. no personal information was transmitted).

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 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 [12]. A closely related study [17] was criticized for not taking into account this "task focus" phenomenon [14]. 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.

#### 4.1.5 Exit Survey

After completing all four study tasks, participants were directed to an online exit survey hosted by SurveyMonkey. The exit survey 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 question 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.

### 4.2 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. 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.

#### 4.2.1 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

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

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.

#### 4.2.2 Effect of Warning Design on Behavior

Our study 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 also 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, users of our multi-page warning visited the library website at an equal rate to users of the FF2 and IE7 warnings. Our single page warning was not significantly different than any 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 ($p < 0.0016$).

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. The

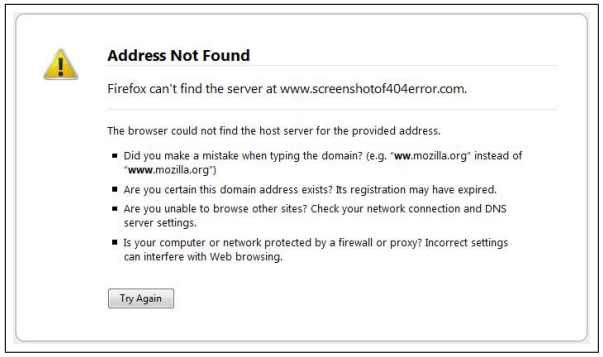



Figure 5: Screenshot of server not found error in FF3.

warnings have very similar layouts and coloring. The yellow Larry icon in the FF3 warning (Figure 3) and the first page of our multi-page (Figure 4(a)) warning is similar to the yellow triangle in Figure 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 warning 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."

#### 4.2.3 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 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.

#### 4.2.4 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* website." 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 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

| 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 6: Behavior in the bank task by reading, understanding, and condition.

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 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 meaning. 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).

#### 4.2.5 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 | Multi |
|---|---|---|---|---|---|
| **Yes** | 8 | 7 | 10 | 4 | 1 |
| **No** | 8 | 11 | 5 | 16 | 16 |
| **Unknown** | 4 | 2 | 5 | 0 | 3 |

Table 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 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. In addition, 5 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 redesigned warnings before. This is dramatic evidence of memory problems, warning confusion, and general confusion with regard to certificate errors. At the same time, it is possible that the novelty of our new warnings

contributed to more participants reading them (and consequently better understanding the risks of ignoring them). None of the participants who viewed our new warnings could have seen them before, while our randomized condition assignments resulted in the two Firefox conditions being assigned 27 participants who were pre-existing Firefox users (68% of 40) and the IE condition being assigned 6 participants who were existing IE users (30% of 20). Thus, it is likely that these 33 participants had already been exposed to the warnings prior to our study, but among our sample population we observed no significant differences in behavior among them and the participants in the IE and FF conditions who were accustomed to using different browsers.

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.

# 5 Discussion

Our warnings somewhat 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, 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 websites. 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 websites like those belonging to banks. If they do become fraud victims, they are unlikely to pinpoint it to their decision to ignore a warning. Thus users' attitudes and beliefs about SSL warnings are likely to undermine their effectiveness [3]. 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.

## 5.1 Limitations

We did not attempt to measure any long term affects of habituation to warnings. Many participants were likely to have previously seen the FF2 and IE7 warnings, while few users were likely to have seen FF3 warnings as that browser was released just before the study began. Our two warnings were new to all participants. We expect users were more likely to ignore the IE7 and FF2 warnings because of habituation, but this is not supported by our data.

Several artifacts of the study design may have caused participants to behave less securely than they normally would. Our study participants knew in advance that they would be using their bank credentials during the study and therefore the most security conscious potential participants may have decided not to perform the study. In addition, the study was performed at and sanctioned by Carnegie Mellon, and therefore participants may have trusted that the study would not put their credentials at risk.

In our study, users were much less likely to heed certificate warnings than in a previous study by Schechter et al. that also examined user responses to the IE7 certificate warning [17]. In our study 90% of participants ignored the IE7 warning while in the Schechter et al. study only 36% of participants who used their own accounts ignored the IE7 warning. We believe the differences may be due to the fact that in the previous study participants were told the study was about online banking, they performed four banking tasks prior to observing the warning, and they were given two other clues that the website might be insecure prior to the display of the warnings. The authors state, "responses to these clues may have been influenced by the presence of prior clues." Furthermore, the previous study was conducted while IE7 was still in beta and thus users were less likely to

have seen the certificate warning before. In addition, our study participants were more technically sophisticated than the previous study's participants.

### 5.2 Explain the Danger

The FF2, IE7, and our single page warnings 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, does 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. In addition, only four users called the library, indicating that our single-page warning would be only a minor nuisance for legitimate websites. That said, we suspect our single page warning would become less effective as users are habituated to it when visiting legitimate websites.

### 5.3 Make it Difficult

The FF3 warning, as discussed at length in Section 4.2.2, 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 safer 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.[12] 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.

[12]See for example: 1) http://hasylab.desy.de/infrastructure/experiment_control/links_and_tutorials/ff3_and_ssl/index_eng.html, 2) http://www.engr.colostate.edu/webmail/, and 3) http://knowledgehub.zeus.com/faqs/2008/02/05/configuring_zxtm_with_firefox_3

### 5.4 Ask a Question

Our multi-page warning, introduced in Section 4.1.2, asks the user a question in order to collect contextual information to allow the browser to better assess the risk of letting the user proceed to the website. This warning suffers from two usability problems: users may answer incorrectly because they are confused and users may knowingly answer incorrectly to get around the warning. In addition, it leaves users susceptible to active attacks such as the finer-grained origins attacks [9]. These problems, plus the fact that the single-page warning was more successful in preventing users from visiting the bank website, lead us to recommend against our multi-page warning as it is currently implemented.

The multi-page warning depends on users correctly answering our question, but only fifteen of the 20 participants answered correctly at the bank website. As discussed in Section 4.2.2, we believe that five participants either knowingly gave the wrong answer in order to reach the destination website without interruption, or they confused the warning with a server unavailable error. However, many users still made mistakes even when answering our question correctly. They behaved no more securely than users of our single-page warning.

Users who answered our question correctly and followed its advice would still be susceptible to finer-grained origins attacks. As brought to our attention by an anonymous reviewer, an attacker with control over the network or DNS may circumvent the multi-page warning by forcing the browser to connect to a website other than the one the user intended. For example, let's say Alice goes to a webmail site (www.mail.com), but an attacker controls the network and wants to steal the password to her online bank (www.bank.com).

When Alice visits mail.com, the attacker sends a response to the Alice that forwards the browser to https://www.bank.com/action.js. Then, the attacker intercepts the connection to the bank with a self-signed certificate, which triggers the warning shown in Figure 4(a). The warning asks her what type of website she is trying to reach and she answers "other" because she believes she is visiting her webmail. Since

Alice answered "other" she is immediately forwarded to action.js. If Alice has an open session with the bank, the attacker steals her bank.com secure cookies with the script.

Even if Alice does not have an open session with the bank, the browser's cache will store the attack script. Let's say in its normal operation the bank site loads its version of action.js after a user logs-in. (If the site loads a different script, then the attacker simply poisons that script instead.) If Alice logs-into www.bank.com in the next year, then the attacker's version of action.js will load instead of the bank's version. As in the attack in the previous paragraph, the script steals her secure cookies. There are many other variations on this attack, but they all rely on Alice answering "what type of website are you trying to visit" based on the site she believes she is visiting instead of the site the attacker sends to her.

Designing an interface to collect contextual information from users without making them susceptible to active attacks such as those outlined above poses a challenge. While we can ask users simple questions about their intentions that they are capable of answering, we must be sure that attackers cannot intervene to mislead users. We may be able to improve the multi-page warning we proposed by asking users another question in certain circumstances. In particular, if the URL of the connecting website is substantially different than the URL the user typed (or clicked on, in the case of a link), then we would show the URL of the connecting website and ask the user if they intended to visit that URL. Unfortunately this is not a complete solution for websites with mixed content, like those using a third-party shopping cart provider. In addition, the usability of such a solution remains untested.

It remains an open research challenge to determine how to leverage contextual information—including user-provided information—in order to assess risks. In particular, an approach is needed that is not vulnerable to confused users, users trying to get around the system, or active attackers. It remains to be seen whether it is feasible to design a robust approach that uses user-provided information. Alternative approaches may leverage contextual information provided by sources other than the user.

### 5.5 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 [10] and Perspectives [20] (and discussed in Section 2), are steps in the right direction. Both systems identify websites likely to be unsafe and use warnings to stop users from proceeding. It would be better to block these unsafe websites entirely. We expect both systems to have extremely low false positive rates, but further evaluation is required to know for sure. Another possible way of identifying unsafe websites is to maintain a list of websites that are verified by a root certificate authority and block websites on the list when the browser receives a self-signed certificate instead.

## 6 Acknowledgements

Thanks to Dhruv Mohindra, Amit Bhan, and Stuart Schechter for their help in the early stages of this project. This work was supported in part by Microsoft Research and by the National Science Foundation under Grants No. 0524189 and 0831428. The first author is supported by a National Defense Science and Engineering Graduate Fellowship.

## References

[1] J. C. Brustoloni and R. Villamarín-Salomón. Improving security decisions with polymorphic and audited dialogs. In *Proceedings of the 3rd symposium on Usable privacy and security*, pages 76–85, New York, NY, USA, 2007. ACM Press.

[2] Certification Authority/Browser Forum. Extended validation SSL certificates, Accessed: July 27, 2007. http://cabforum.org/.

[3] L. F. Cranor. A framework for reasoning about the human in the loop. In *Proceedings of the 1st Conference on Usability, Psychology, and Security*, pages 1–15, Berkeley, CA, USA, 2008. USENIX Association.

[4] R. Dhamija, J. D. Tygar, and M. Hearst. Why phishing works. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 581–590, New York, NY, USA, 2006. ACM.

[5] I. E-Soft. SSL server survey, February 1, 2007. http://www.securityspace.com/s_survey/sdata/200701/certca.html.

[6] S. Egelman, L. F. Cranor, and J. Hong. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1065–1074, New York, NY, USA, 2008. ACM.

[7] 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 SIGCHI Conference on in Computing Systems*, Seattle, WA, March 31 - April 4, 2001. ACM.

[8] B. Friedman, D. Hurley, D. C. Howe, E. Felten, and H. Nissenbaum. Users' conceptions of web security: a comparative study. In *Extended Abstracts on Human Factors in Computing Systems*, pages 746–747, New York, NY, USA, 2002. ACM.

[9] C. Jackson and A. Barth. Beware of finer-grained origins. In *Proceedings of the Web 2.0 Security and Privacy Workshop*, 2008.

[10] C. Jackson and A. Barth. ForceHTTPS: protecting high-security web sites from network attacks. In *Proceeding of the 17th International World Wide Web Conference*, pages 525–534, New York, NY, USA, 2008. ACM.

[11] C. Jackson, D. R. Simon, D. S. Tan, and A. Barth. An evaluation of extended validation and picture-in-picture phishing attacks. In *Proceeding of the 1st International Workshop on Usable Security*, pages 281–293, Berlin / Heidelberg, Germany, February 2007. Springer.

[12] S. Milgram. *Obedience to Authority: An Experimental View*. Harpercollins, 1974.

[13] J. Nightingale. SSL information wants to be free, January 2009. http://blog.johnath.com/2009/01/21/ssl-information-wants-to-be-free/.

[14] 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/.

[15] 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.

[16] 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.

[17] S. E. Schechter, R. Dhamija, A. Ozment, and I. Fischer. The emperor's new security indicators. In *Proceedings of the 2007 IEEE Symposium on Security and Privacy*, pages 51–65, Washington, DC, USA, 2007. IEEE Computer Society.

[18] J. Sobey, R. Biddle, P. C. van Oorschot, and A. S. Patrick. Exploring user reactions to new browser cues for extended validation certificates. In *Proceedings of the 13th European Symposium on Research in Computer Security*, pages 411–427, 2008.

[19] 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.

[20] D. Wendlandt, D. G. Andersen, and A. Perrig. Perspectives: Improving SSH-style host authentication with multi-path probing. In *Proceedings of the 2008 USENIX Annual Technical Conference*, Berkeley, CA, USA, June 2008. USENIX Association.

[21] 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.

[22] M. Wogalter. Purpose and scope of warnings. In M. Wogalter, editor, *Handbook of Warnings*, pages 3–9. Lawrence Erlbaum Associates, Mahway, NJ, USA, 2006.

[23] M. Wu, R. C. Miller, and S. L. Garfinkel. Do security toolbars actually prevent phishing attacks? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 601–610, New York, NY, USA, 2006. ACM.

[24] H. Xia and J. C. Brustoloni. Hardening web browsers against man-in-the-middle and eavesdropping attacks. In *Proceedings of the 14th International World Wide Web Conference*, pages 489–498, New York, NY, USA, 2005. ACM.