Experimenting At Scale With Google Chrome's SSL Warning

Adrienne Porter Felt
Robert W. Reeder
Google Inc.
felt, rreeder@google.com

Hazim Almuhimedi
Carnegie Mellon University
hazim@cs.cmu.edu

Sunny Consolvo
Google Inc.
sconsolvo@google.com

ABSTRACT

Web browsers show HTTPS authentication warnings (i.e., SSL warnings) when the integrity and confidentiality of users' interactions with websites are at risk. Our goal in this work is to decrease the number of users who click through the Google Chrome SSL warning. Prior research showed that the Mozilla Firefox SSL warning has a much lower click-through rate (CTR) than Chrome. We investigate several factors that could be responsible: the use of imagery, extra steps before the user can proceed, and style choices. To test these factors, we ran six experimental SSL warnings in Google Chrome 29 and measured 130,754 impressions.

Author Keywords

Browser security warnings; SSL warnings; interruptive warnings; active warnings; interstitials

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation (e.g. HCI): User Interfaces; K.6.5 Management of Computing and Information Systems: Security and Protection

INTRODUCTION

Web users rely on SSL for the privacy and security of their data. For journalists and dissidents, SSL can be the difference between safety and physical harm. Browsers show SSL warnings when they cannot establish a well-authenticated HTTPS connection to a website. When these warnings appear, it is up to the user to decide whether to proceed.

Our goal is to decrease the number of users who click through (i.e., ignore) Google Chrome's SSL warnings. Clicking through an SSL warning can be a safe choice if the user is confident that the warning is due to a benign server misconfiguration. However, it is often difficult or impossible to differentiate between server misconfigurations and attacks. Separate efforts are needed to improve the precision of SSL warnings, but we focus on nudging users in the direction of a lower CTR. We aim for a lower CTR because (a) it's safer to err on the side of caution, and (b) we hope that low CTRs will encourage developers to adopt valid SSL certificates.

CHI 2014, April 26 - May 01 2014, Toronto, ON, Canada

http://dx.doi.org/10.1145/2556288.2557292

Usable security researchers have studied web browser security warnings for years [4, 8, 2]. However, the difficulty of creating ecologically valid laboratory studies of warnings has impeded warning research. Participants may behave unnaturally in a laboratory setting [7]. Even when some idiosyncracies of laboratory studies are mitigated, experimenters still have to use contrived designs to direct participants toward sites where warnings will appear.

The most natural way to study SSL warnings is to measure reactions to real warnings on users' computers. We measured user reactions to experimental warnings encountered during everyday browsing in Google Chrome. In this paper, we present findings from 130,754 warning impressions. We implemented six experimental warnings in Google Chrome 29 that are designed to test several hypotheses about how users respond to warning design manipulations.

Akhawe and Felt showed that Firefox's SSL warning has a considerably lower CTR than Chrome's (33% vs. 70%) [1]. We tested the hypothesis that it is the warning's design — rather than the characteristics of Firefox or its user population — that leads to Firefox's lower CTR. We further tested whether any design advantages of the Firefox warning were due to: its requirement of an extra step to proceed through the warning; its distinctive, non-commercial styling; or its use of a human image with its gaze directed at the user.

Contributions. We make the following contributions:

- We show that warning design can drive users towards safer decisions. Design accounted for between a third and half of the difference in CTRs between Chrome and Firefox.
- Warning design did not account for the remaining difference between browsers. This means that other factors influence the CTR.
- Several design variations, such as images of watching eyes, had little to no effect on behavior.
- To our knowledge, we are the first to publish a field experiment on the effects of browser warning design under realistic conditions.

METHODOLOGY

We deployed six experimental SSL warnings and one matched control as part of Google Chrome 29. We measured user reactions to the default Chrome SSL warning (Condition 1), three versions of the Chrome SSL warning with new images (Conditions 2-4), and a replica of the Firefox SSL warning with two variants (Conditions 5-7).

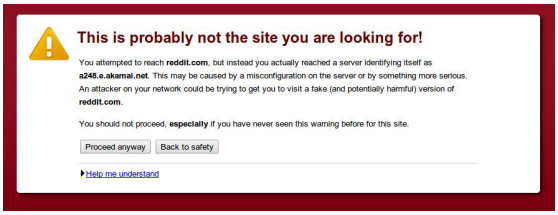

Figure 1. The default Chrome SSL warning (Condition 1).

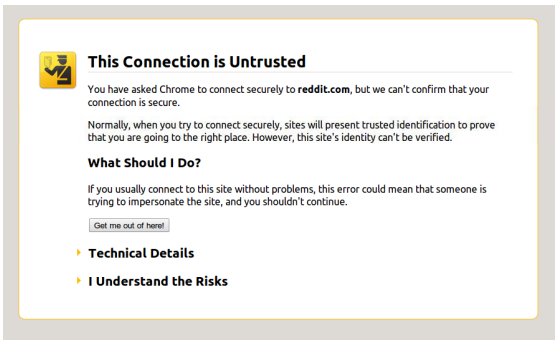

Figure 2. The mock Firefox SSL warning (Condition 5).

Hypotheses and Conditions

Firefox Warning Appearance

Hypothesis: The visual design of the Firefox SSL warning is the reason for the lower CTR in Firefox.

Akhawe and Felt found that Firefox's SSL warning has a CTR of 33% whereas Google Chrome has a CTR of 70% [1]. To test the impact of visual design on the CTR, we implemented a replica of the Mozilla Firefox SSL warning in Google Chrome.[1] Figure 1 shows the default Chrome SSL warning (Condition 1), and Figure 2 shows the mock Firefox SSL warning (Condition 5). Demographics, browsing habits, and other non-appearance factors are held constant because they were both tested in Google Chrome.

Our mock Firefox warning is identical to the actual Firefox warning in all ways but two. First, we replaced the name "Firefox" with "Chrome" in the warning text. Second, proceeding through the actual Firefox warning yields a secondary pop-up dialog that asks whether the browser should permanently remember the user's decision to proceed. Google Chrome did not support this feature at the time of this experiment, so there is no secondary dialog.

Steps to Proceed Past the Warning

Hypothesis: An extra step will *decrease* the CTR.

Some designers add extra steps to warnings with the intention of reducing the CTR. For example, Firefox users need to take three steps to proceed through the Firefox SSL warning: (1) click on "I Understand the Risks," (2) click on the (now unhidden) button to proceed, (3) click through a final pop-up dialog that appears in a separate window.

Sunshine et al. showed that the extra steps in the Firefox SSL warning make it difficult for users to proceed through the warning [8]. However, they conjectured that this would only work until users learn how to work around it. Akhawe and Felt studied the Firefox SSL warning and found that the third step discourages 15% of users from proceeding further, but they did not collect data on the earlier steps [1].

[1]With approval from the author of the Firefox warning.

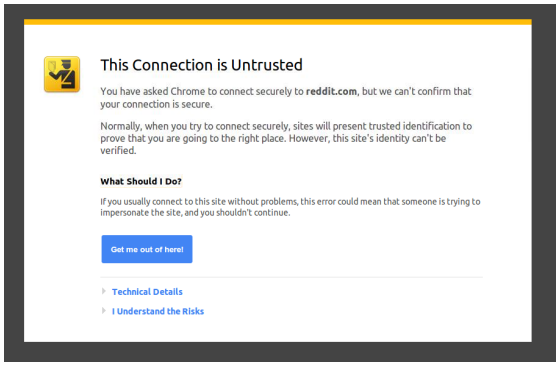

Figure 3. The Firefox SSL warning with Google styling (Condition 7).

To bypass Conditions 5, 6, and 7, participants must: (1) click on "I Understand the Risks" to reveal the proceed button, (2) click the proceed button. We recorded how many participants clicked on both steps so that we could see how many participants changed their minds due to the extra step.

Corporate Style Guidelines

Hypothesis: Applying corporate style guidelines to a warning will *increase* the CTR.

We hypothesize that warnings that resemble corporate products will have higher CTRs because they do not stand out as unusual. To test this, we created a Google-styled version of the Firefox SSL warning. Condition 5 is a faithful replica of the Firefox SSL warning, with a gray palette and unstyled buttons and links (Figure 2). A Google designer created another version by applying Google's corporate style guidelines to the warning (Condition 7). Condition 7 uses Google's palette, Google-styled buttons, and Google-styled links (Figure 3). We kept the text and layout constant between the two versions. Although Condition 7 could have been made to look more like a Google product if we had altered the text and layout, we wanted to control for these factors.

Images of Watching People

Hypothesis: Including an image of a human in a warning will *decrease* the CTR.

Studies have found that people behave in a more socially conscious manner when they are near images of watching eyes [5, 6]. Detecting a human face in an image activates the "social brain," which encourages pro-social and cooperative behavior [6]. We hypothesize that this physiological effect would lead to a lower warning CTR.

The Firefox warning (Condition 5, Figure 2) contains a black image of a human figure on a yellow-orange background. Although this figure does not have eyes or a face, it should still create the sensation of being watched because its posture indicates that it is looking at the viewer [3]. For comparison, Condition 6 is the same warning without the image.

Figure 4. The three images used in Conditions 2-4.

We added two images of human faces to the Chrome SSL warning: a policeman (Condition 2) and a criminal (Condition 3). Their eyes stare directly at the viewer. The images are drawings, which prior work has shown to be sufficient to activate the social brain [5]. For comparison, Condition 4 includes a red traffic light; the traffic light conveys the same "stop" message, but without a human face. Figure 4 shows the three images, which were the same height as the first paragraph of the Chrome warning (Figure 1).

Field Study Deployment

We modified Google Chrome 29 to include our experimental versions of the warnings. The first time a Google Chrome 29 client begins to load an SSL warning, our field trial code pseudorandomly assigns the client to a condition and loads the appropriate version of the warning. For each condition there was a 1.4% chance that the client would be assigned to it. A given client could be assigned to only one condition. The remaining 90.2% of the population received the default behavior and was not part of the study.

Google Chrome's opt-in metrics allow us to measure reactions to security warnings. During installation, Chrome users are asked whether they would like to send "crash reports and statistics" to Google. If they choose to participate, Chrome periodically sends statistical reports to Google. Each report includes whether the user has recently seen or clicked through an SSL warning, and this data is tagged with the appropriate condition. This lets us correlate CTRs with our experimental conditions. The reports are pseudonymous and, once stored, cannot be traced back to the sending client.

Our study ran from August 22 to 31, 2013. We report data from Google Chrome 29 (stable). Our data is from English (U.S.) clients on Windows, Mac, Chrome OS, and Linux.

Experimental Ethics

We relied on Google Chrome's opt-in metrics to measure click-through rates. We did not collect any sensitive or personal information about participants (e.g., no browsing history). We followed our internal review processes for field trial design quality and privacy.

One concern was that our experiment could increase the CTR, thereby putting users at greater risk. The study was first deployed on a small scale to developer versions of Chrome in May 2013, and we monitored the CTRs of the conditions. If any of the conditions had yielded adverse effects, we would have halted those conditions; however, they did not.

Limitations

Our sample is limited to participants in Google Chrome's metrics program. Since this is an opt-in program, it is possible that there is selection bias in our sample. However, even if they are not representative of the whole population, they still constitute a notable minority.

#	Condition	CTR	N
1	Control (default Chrome warning)	67.9%	17,479
2	Chrome warning with policeman	68.9%	17,977
3	Chrome warning with criminal	66.5%	18,049
4	Chrome warning with traffic light	68.8%	18,084
5	Mock Firefox	56.1%	20,023
6	Mock Firefox, no image	55.9%	19,297
7	Mock Firefox with corporate styling	55.8%	19,845

Table 1. Click-through rates and sample size for conditions.

Although we restricted each client to receiving only one condition, it is possible that participants with multiple computers experienced multiple conditions.

RESULTS AND IMPLICATIONS

We observed CTRs ranging from 55.8% to 68.9% for the six conditions and control. Table 1 contains an overview of the conditions and CTRs. In the following section, we correct for multiple testing by lowering our overall $\alpha = 0.05$ to $\alpha = 0.0083$ using Bonferroni's adjustment.

Firefox Warning Appearance

We find that visual appearance accounts for between a third and half of the 37-point (70%-33%) difference between Chrome' and Firefox's CTRs. We calculate this as follows:

- Participants clicked through 67.9% of default Chrome warnings (Condition 1) and 56.1% of mock Firefox warnings (Condition 5). Since all other factors were held constant, differences in the warnings' appearances are responsible for 12 of 37 points.

- Firefox users see a pop-up confirmation dialog after experiencing the real warning. 15% of the time that users see this dialog, they turn back [1]. If we were to implement this dialog in Chrome, it might have reduced the CTR by another 15%. This would make the warning as a whole responsible for an additional 8 points (15% $\times$ 56.1%).

Novelty could potentially bias participants' responses to the mock Firefox warning. Participants might have been startled or intrigued by an unfamiliar warning, leading to a lower CTR. However, the overall CTR remained steady for the duration of the study, and the CTR for participants with repeat impressions did not vary. Either ten days is insufficient for novelty to wear off, or novelty did not contribute to the CTR.

The control condition yielded a CTR of 67.9%, whereas Akhawe and Felt previously reported a CTR of 70% for Chrome [1]. A small amount of the difference could be attributed to fluctuation over time.

We therefore estimate that the design of the warning and pop-up dialog together account for between 12 and 20 points (i.e., 32% to 54%) of the difference between the two browsers' CTRs. This demonstrates that design can influence users' security decisions. The remaining difference must be due to

other factors. Different demographics[2] might have different risk tolerances or preferences. Other aspects of the user experience might also change how users perceive warnings.

Steps to Proceed Past the Warning

For Conditions 5, 6, and 7, participants had to click twice to proceed past the warning. The second step did not serve as a meaningful deterrent: for all three conditions, 98% of participants who performed the first step also completed the second step. This demonstrates that the addition of a very simple extra step may not have a notable effect on the CTR. However, Akhawe and Felt reported that only 85% of users clicked through Firefox's third step (a pop-up dialog with more technical information), which means the third step is a bigger deterrent [1]. Combined with our finding, this suggests that the effectiveness of an extra step may depend on its complexity.

Corporate Style Guidelines

Applying Google's corporate style guidelines to the mock Firefox warning did not increase the CTR. The Google-styled version of the warning (Condition 7) performed slightly better than the unmodified mock Firefox warning (Condition 5), which is the opposite of what we predicted. However, the difference is very small (56.1% vs. 55.8%). We interpret this result to mean that tweaks to the color and style – e.g., updating an old warning with a newer style guide – may not have an effect on the CTR.

We held the layout and wording constant between Conditions 5 and 7 to avoid potential confounds. It is possible that changing the layout and wording to look more like a commercial product would yield the anticipated effect.

Images of Watching People

The brain's social response to human images is instinctive, and it should occur for even a hint of a human face [6, 5]. If the feeling of being watched were to influence how users react to warnings, all of the conditions with human images should have lower CTRs. However, we did not find this.

- Removing the human figure from the mock Firefox warning did not have an effect (56.1% vs. 55.9%) [1-tail z-test of proportions, $p = .3485$].
- The policeman (Condition 2) performs slightly worse than the imageless default warning (Condition 1): 67.9% vs. 68.9%, which was the opposite of our hypothesis.
- The criminal (Condition 3) had a lower CTR than the control (Condition 1) by a statistically significant amount [1-tail z-test of proportions, $p = 0.0025$], but the effect size is very small (67.9% vs. 66.5%). It also had a lower CTR than the red traffic light, which served as a secondary control [1-tail z-test of proportions, $p < 0.0001$].

Although ignoring an SSL warning can have social implications (e.g., leaking others' social media posts), this may not occur to participants when they are viewing warnings. Thus, triggering the social portion of the brain may not influence their decisions. The criminal may have yielded a very slight improvement through a different mechanism: fear arousal.

Other Design Differences

We found that the design of the Mozilla Firefox warning without the pop-up accounts for a third of the difference between the two browsers. What makes it more effective? We have ruled out the image of a human, the first additional step, and the styling as the cause. We therefore hypothesize that the Firefox warning's text, layout, and/or default button choice are responsible. The Firefox warning appears to follow warning design guidelines from prior work. The warning avoids technical jargon, identifies ways to mitigate the risk under "What Should I Do?" [9], hides technical details by default [4], and has a clear default choice [4, 2].

[2]http://elie.im/blog/web/survey-internet-explorer-users-are-older-chrome-seduces-youth/

ACKNOWLEDGEMENTS

We thank Johnathan Nightingale for allowing us to replicate Firefox's warnings; Roberto Ortiz and Sebastien Gabriel for designing the new SSL artwork; and Melissa Bateman and Ross Anderson for discussing the use of human images.

REFERENCES

1. Akhawe, D., and Felt, A. P. Alice in Warningland: A Large-Scale Field Study of Browser Security Warning Effectiveness. In *USENIX Security Symposium* (2013).
2. Egelman, S., Cranor, L. F., and Hong, J. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In *Proceedings of CHI* (2008).
3. Emery, N. The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews 24* (2000).
4. Nodder, C. Users and trust: A Microsoft case study. *Security and Usability: Designing Secure Systems that People Can Use* (2005), 589–606.
5. Rigdon, M., Ishii, K., Watabe, M., and Kitayama, S. Minimal social cues in the dictator game. *Journal of Economic Psychology 30* (June 2009).
6. Senju, A., and Johnson, M. H. The eye contact effect: mechanisms and development. *Trends in Cognitive Science* (March 2009).
7. Sotirakopoulos, A., Hawkey, K., and Beznosov, K. On the Challenges in Usable Security Lab Studies: Lessons Learned from Replicating a Study on SSL Warnings. In *Proceedings of SOUPS* (2011).
8. Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., and Cranor, L. F. Crying Wolf: An Empirical Study of SSL Warning Effectiveness. In *USENIX Security Symposium* (2009).
9. Wogalter, M. S., Conzola, V. C., and Smith-Jackson, T. L. Research-based guidelines for warning design and evaluation. *Applied Ergonomics 33*, 3 (2002).