The Web's Identity Crisis: Understanding the Effectiveness of Website Identity Indicators

Christopher Thompson, Martin Shelton, Emily Stark,
Maximilian Walker, Emily Schechter, Adrienne Porter Felt
Google

Abstract

Users must understand the identity of the website that they are visiting in order to make trust decisions. Web browsers indicate website identity via URLs and HTTPS certificates, but users must understand and act on these indicators for them to be effective. In this paper, we explore how browser identity indicators affect user behavior and understanding. First, we present a large-scale field experiment measuring the effects of the HTTPS Extended Validation (EV) certificate UI on user behavior. Our experiment is many orders of magnitude larger than any prior study of EV indicators, and it is the first to examine the EV indicator in a naturalistic scenario. We find that most metrics of user behavior are unaffected by its removal, providing evidence that the EV indicator adds little value in its current form. Second, we conduct three experimental design surveys to understand how users perceive UI variations in identity indicators for login pages, looking at EV UI in Chrome and Safari and URL formatting designs in Chrome. In 14 iterations on browsers' EV and URL formats, no intervention significantly impacted users' understanding of the security or identity of login pages. Informed by our experimental results, we provide recommendations to build more effective website identity mechanisms.

1 Introduction

To use the web safely, users must be able to understand the identity of the website that they are visiting. Without understanding a website's identity, users cannot make an informed decision about whether to provide it with their personal information or trust its content. Such misunderstandings result in common attacks like phishing and social engineering [36].

Web browsers use two mechanisms to communicate website identity to users. The first is the URL displayed in the browser address bar, along with a padlock icon to indicate an authenticated connection. For example, before a user types their Google password into a webpage, the user should verify that the domain in the browser address bar is "google.com" and that the padlock icon is present. Second, some HTTPS connections are authenticated with an Extended Validation (EV) certificate. An EV certificate associates a website with a legal entity, whose name and jurisdiction is typically displayed alongside the URL in the address bar (Figure 1). Users can check the EV indicator to verify that a website is associated with an established legal entity that they trust.

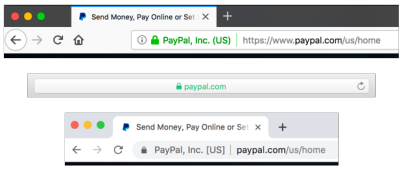


Figure 1: Examples of EV certificate UI in different web browsers (from top to bottom: Firefox, Safari, and Chrome).

Prior work suggests that neither URLs nor EV indicators work very well as indicators of website identity. Many users do not look at the URL even when primed to try to identify fraudulent sites [12]. EV indicators also do not help users identify fraudulent sites [24]. Even if users did notice the EV indicator, EV certificates can contain misleading information that limits their usefulness [10, 11].

However, much of the work that studies the effectiveness of website identity indicators is dated, testing browser UIs that are no longer in common use. For example, most EV indicator research is ten years old and studied the first browser EV UI from Internet Explorer 7. Browser security indicators and the web security landscape have changed dramatically since these studies were conducted, with widespread adoption of HTTPS [15] and constant evolution of browser UI [16]. Further, prior work does not examine how users react to website identity indicators in the wild.

In this paper, we examine the effectiveness of browser identity indicators – URLs and EV UI – from several angles. We focus primarily on how users react to the EV indicator, since it is designed to provide human-meaningful identity information, and we also investigate whether browser UIs can be tweaked to make the URL more human-meaningful. Our

goal is to study the effectiveness of modern browser identity indicators at a much larger scale than previous work.

First, we analyzed a large-scale field experiment (Section 3) from the Google Chrome web browser. We examined a suite of user behavior measurements with and without the EV indicator present on websites that serve EV certificates. This experiment simulates a situation in which a user visits an attack website that mimics a victim website exactly but does not possess an EV certificate for the victim site. We find little evidence that the absence of the EV indicator affects how users interact with the site. We do find, however, that the EV indicator itself draws clicks.

Second, we conducted a series of survey experiments (Section 4) to investigate follow-up questions about EV UI that we could not answer with field data. Our surveys, with over 1,800 total participants from the U.S. and U.K., sought to answer two questions: (1) Would a recent proof-of-concept attack on EV certificates [11] be effective on real users? (2) How do users react to other browsers' EV UIs? In these surveys, we find no evidence that the EV UI in either Chrome or Safari impact how comfortable users feel when logging into a webpage.

Finally, having found little evidence that EV indicators influence user behavior, we consider whether URLs can be more effective identity indicators (Section 5). We surveyed over 1,000 users to assess reactions to different variations on Chrome's URL display. Each variation was designed to draw users' attention to the domain name, in hopes that they would notice that the webpage was a phishing site. We found no significant differences among any of the variations, leading us to believe that a more radical redesign is necessary for URLs to effectively communicate website identity to users.

Our results, along with the body of existing work, suggest that modern browser identity indicators are not effective. We use these results to provide recommendations for building better website identity mechanisms. Based on promising results from prior research [7, 14], we argue that negative, active indicators (such as full-page warnings) are a more promising avenue than passive, positive indicators (such as the padlock icon or EV indicator in the address bar). We further recommend that user research should be incorporated into the design phase for future browser identity indicators.

2 Background

2.1 URLs and website identity

The fundamental identity indicator of the web is the URL. All major web browsers display the URL of the page in order to convey the website's identity. Figure 2 shows how different web browsers display URLs to users.

2.1.1 HTTPS and certificates

In the URL bar, the presence of the "https://" scheme and/or a padlock icon indicate that the identity of the site has been verified through a cryptographic certificate. Most websites use *domain-validated (DV) certificates*. A website owner can obtain a DV certificate by proving control over a domain to a certificate authority (CA) [1]. DV certificates can be obtained easily and inexpensively from a number of CAs.

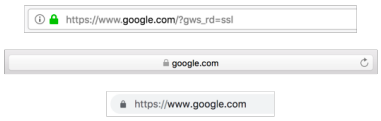


Figure 2: Examples of different browser URL displays (from top to bottom: Firefox 64, Safari 12.0.3, and Chrome 72).

2.1.2 Registrable domain

In most situations, a user who is making a security decision should pay attention to the *registrable domain* [6] rather than the full URL. The registrable domain is composed of the high-level domain name suffix under which internet users can register names [5], plus the DNS label immediately preceding it. For example, "google.com" is the registrable domain in the URL "https://accounts.google.com/", and "google.co.uk" in "https://accounts.google.co.uk/".

The registrable domain is typically the identity indicator of interest because when an organization controls a registrable domain, the same organization can typically control all the subdomains and paths within that domain.

2.2 Extended Validation (EV) certificates

EV certificates are a type of HTTPS certificate in which a domain owner undergoes additional validation with a CA to tie their domain to a legal entity. Most major web browsers display the legal entity name and jurisdiction in the URL bar alongside the URL, as shown in Figure 1. Notably, Safari recently stopped displaying the legal entity name and simply colors the domain green when an EV certificate is present.[1]

Section 2 of the CA/Browser Forum guidelines for EV certificates [3] specifically states that one primary purpose of EV certificates is to enable a browser to inform the user about the specific legal identity of the business with which they are interacting when using a website. A secondary purpose given by the guidelines is that EV certificates can be considered to combat phishing and other malicious web activity.[2] If an attack website is impersonating a victim business, the attack site is not supposed to be able to obtain an EV certificate for the victim business. A user visiting the attack website might notice that there is no EV indicator for the victim business and thereby conclude that the website is not the legitimate website for the victim business.

EV certificates are typically more cumbersome to obtain than DV certificates. Domain owners pay a premium for

[1] Apple did not mention this change in their release notes, however it was discussed on Twitter [13, 30] and technical blogs [23].

[2] Microsoft [18] and Mozilla [21] gave similar motivations for their introduction of browser EV UI.

EV certificates and undergo a days- or weeks-long validation process.

2.2.1 Weaknesses of EV certificates

EV certificates suffer from a number of security and usability weaknesses. These weaknesses have led to a vigorous debate in the security community about whether browsers should continue to display EV certificate UI [22, 23, 32]. Our work seeks to inform this debate with up-to-date, large-scale, in situ data about how users react to EV indicators.

Malicious EV certificates. EV is not intended to verify that the holder of the certificate is law-abiding, trustworthy, or safe [3]. Researchers have shown that EV certificates can be obtained for misleading names that could be useful in an attack. We describe these attacks in Section 7.1.2.

Usability issues. A body of work from the mid-2000s indicates that EV certificates are not an effective phishing defense because users do not pay attention to the EV UI in web browsers. We survey this work in Section 7.1.1.

Further, the legal entity names in EV certificates are not always intuitive or understandable, because the legal entity does not always match the company's user-visible brand. For example, the personal finance management site `mint.com` has a legal entity name of "Intuit Inc."

3 EV field experiment

To understand whether browser EV UI has an effect on user behavior in the wild, we analyze data from a large-scale field experiment. In this experiment, the EV UI was disabled for a subset of Google Chrome users. We compare a variety of metrics representing users' interactions with EV websites in the experimental and control groups. We do not find evidence that the EV UI impacts user behavior significantly for most metrics. The exception is that users who see the EV UI are more likely to open and interact with the Page Info bubble, which is anchored to the connection security indicator chip (Figure 3). Additionally, we examined the effects of removing the EV UI on a set of 20 top EV sites. We found a small negative impact on navigations to one of these sites.

3.1 Methodology

3.1.1 Dataset

We analyze data from Chrome's user metrics program. Chrome collects metrics in the form of enums, booleans, counts, and times. Our dataset comes from the Stable channel, which has the largest set of users and is the default installation release channel. Stable channel is considered the most representative for measurement and experimentation purposes.

Chrome metrics reports are pseudonymous, containing client information such as the operating system and country, but no personal information (e.g., age or gender). The user metrics program is enabled by default for consumer installs. Users may opt out during installation or in browser settings.

A subset of metrics are keyed by the URL on which the metric is recorded. URLs are only provided for users who have also opted to sync their browsing data with Google servers [20]. We use these URL-keyed metrics to check for changes in user behavior on specific well-known sites with EV certificates. We analyze a subset of our metrics on each of the top 20 EV sites, as visited by Chrome users during our experimental period. We report these results with the domains blinded.

Our dataset includes metrics collected from January 15, 2019 to January 28, 2019. Chrome 71 was fully rolled out to the Stable channel during this time period.[3]

3.1.2 Experimentation framework

Chrome contains an experiment framework in which users can be randomly assigned to experiment groups. Metrics reports are then tagged with the user's group. In our dataset, 1% of Stable channel users were assigned to an experimental group in which the EV certificate UI was disabled, and 1% to a control group. In the experimental group, users who visited EV sites saw a padlock but no additional HTTPS UI.

3.1.3 Metrics

In our study, we analyzed a set of user behaviors that we hypothesized might be affected by a user's perception of the security and identity of a site.

Our selection of metrics is informed by a review of related work (Section 7). We sought to measure behaviors in situ that previous work measured via lab studies or surveys. Below we describe each metric and why we included it.

- **Navigations.** Do users navigate to different sites, or navigate away from EV sites more in the experimental group? We measure:
  - The number of navigations to EV pages, normalized by the total number of navigations.
  - The median time spent on each page visit (not including time spent in the background).[4]
  - The number of times that users left EV pages, by closing the tab, using Back/Forward functionality, or reloading. We normalize by the total number of navigations to EV pages.

  We are interested in navigations because prior work on browser security indicators surveyed users on whether they would leave the page if a particular security indicator appeared in their browser [16].
- **Form submissions.** A metric is recorded when a user submits an HTML form. We consider the number of

[3] During this period, our dataset included millions of clients in each group for our main analysis. For our per-origin analysis, we had tens of thousands of clients on average in each group for each origin.

[4] Due to a bug in our initial data collection for this metric, we instead use data from May 15–28 (after a fixed version reached Chrome's Stable channel) for this metric.

form submissions that occur on pages with EV certificates, normalized by the total number of navigations to EV pages. Previous studies have surveyed users on their willingness to enter login [31] or credit card [16] details, both of which typically involve submitting a form.

- **Autofill interactions.** Chrome's autofill feature saves credit card details that the user enters and provides suggestions when users fill out payment forms. We analyze the number of times a suggestion was selected normalized by the number of times a suggestion was shown. As with form submissions, these metrics provide insight into whether users in the experimental group are less comfortable providing credit card details to the page.
- **Page Info interactions.** Chrome's Page Info bubble is the dialog that appears when a user clicks on the main connection security indicator in the address bar (Figure 3). A metric is recorded every time a user opens the Page Info bubble and every time they use its functionality (e.g., opening the certificate details dialog or inspecting cookies). We analyze Page Info behavior because previous lab studies have examined users' interactions with the equivalent dialog in other browsers [33]. We normalize the number of times the Page Info bubble was opened by the total number of navigations to EV pages. We normalize the number of different actions within the Page Info bubble by the total number of times the Page Info bubble was opened.
- **Downloads.** Downloading a file, particularly an executable, may represent a trust decision for users. We record the number of downloads initiated from EV pages, normalized by the total number of EV navigations.
- **Site Engagement.** Chrome records an aggregate metric called Site Engagement (SE) that approximately measures how much active time a user spends on a site.[5] Each web origin receives a SE score between 0 and 100. It goes up as a user clicks, scrolls, performs keypresses, or plays media on a site, and decays over time as a user does not interact with the site. We compare SE scores with and without the EV UI to see if there might be effects related to user engagement that are not captured by our other metrics. We analyze this metric on a per-site basis. For each user on each origin, we compute the average SE score per visit as well as the average change in SE score over each visit.

3.1.4 Analysis

To see if there are statistically significant effects on any of our metrics of user behavior between our control group and our experimental group, we perform a Welch's t-test for unequal sample variances (as our sample sizes and variances

[5] `https://www.chromium.org/developers/design-documents/site-engagement`

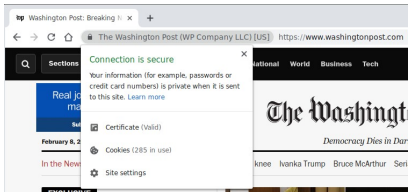


Figure 3: The Page Info bubble in Chrome.

are not guaranteed to be equal between our treatment and control groups) for each metric of interest.

For each metric, we report the difference between the experimental group and the control group, the 95% confidence interval for the difference, and the p-value of the t-test.

3.1.5 Limitations

Incompletely capturing user reactions. It is possible that users in the experimental group reacted differently than the control group in ways that we did not measure. For example, perhaps when the EV UI is disabled, users use a throwaway password or deny all permission prompts. Though we cannot feasibly measure all such user reactions, we feel that our study still provides value by (a) measuring a wide variety of user behaviors that one could reasonably expect to be influenced by a security or identity indicator, and (b) studying user behaviors in a naturalistic scenario that have previously been studied only in labs or surveys.

Limited insight into per-site effects. Some, but not all, of our metrics are keyed by URL (Section 3.1.1). For metrics which are not URL-keyed, we can draw conclusions only about user behavior in aggregate over all sites. It is possible that the absence of the EV indicator influenced these metrics on particular sites but did not have a significant effect when aggregated over all EV sites.

Incomplete simulation of attack scenarios. Our study analyzes whether users react to a missing EV indicator on a website that does not otherwise look suspicious. This simulates an attack scenario in which an attack website mimics a victim website exactly except for the EV certificate. This attack scenario could arise if an attacker obtains a homograph domain (one that looks nearly or exactly identical to a victim domain)[6] or an ordinary domain-validated certificate for the victim site. However, EV certificates might significantly influence user behavior in other attack scenarios that we did not study. For example, consider a website that spoofs `paypal.com` but is hosted at an obviously incorrect URL. In this scenario, a missing EV indicator might prompt a user to inspect the URL and thereby detect the attack. While our experiment does not cover such attack scenarios, we feel that our experiment's attack scenario is of particular interest: the purpose of EV indicators are to provide human-meaningful identity information on the premise that other signals, such

[6] `https://en.wikipedia.org/wiki/IDN_homograph_attack`

as the URL, are not sufficient tools for identifying websites. We explore an additional attack scenario of practical interest via a survey experiment, as described in Section 4.

Dataset selection. Our dataset does not come from a truly random sample of Chrome's user population: users can opt out of the user metrics program, and users must specifically opt in to browsing data syncing to report URLs (Section 3.1.1). However, we still believe this data is valuable in light of its scale and naturalistic observation.

3.1.6 Ethical considerations

Although our institution is not subject to IRB approval, the EV experiment went through an internal review process before launching, including security and privacy reviews. As discussed in Section 3.1.1, Chrome metrics reports are pseudonymous [20].

The experiment rolled out in several release channels before Stable, per Chrome's usual release process. The experiment was monitored as it rolled out, and had any problems been detected, it could have been disabled at any time.

For users in the experimental group, the Chrome developer console contained a message explaining the experiment. This message was intended to inform site owners why their EV certificate UI might not be showing.[7]

Changes to browser security and identity indicators come with the risk that users feel safer on malicious sites and take actions that they wouldn't otherwise take (for example, a user might enter credit card details on a scam site because the UI change made them believe it was safe). Our approach is similar to other field studies on browser security UI, such as exploring new security indicator icons [16], and more conservative than default feature rollouts in Chrome, as the experiment targeted only a small percentage of users and could have been disabled at any time had there been unexpected effects indicating that the experiment put users at risk. In this case, we expected the experiment to, at most, make users act more cautiously on legitimate sites, since we were only modifying positive security UI (compared to, for example, prior work experimenting with full page connection security warnings [14]).

We note that Brave Browser, which is based on the Chromium project, has opted in to not showing the EV UI using our experimental feature [2]. Brave's previous implementation (based on Muon) also intentionally did not show any EV UI [4]. Our dataset only includes data from official Chrome clients.

[7]The developer console is a default-hidden UI intended for web developers, where many technical warnings about the page are printed (e.g., identifying specific mixed content subresources, or the use of deprecated APIs). In the Stable channel, the console was opened by 2% of clients over the 14-day period of our study. We believe that this indicates that the console warning would not be a potential source of priming for a vast majority of participants. We did not see a significant difference in how often the console was opened (normalized by page loads) between our experimental and control groups (p=0.57, 95% CI: [-0.000052,+0.000028]).

3.2 Results

3.2.1 Summary

We did not see any significant differences in user behavior in our navigation or on-page metrics between our experimental group and our control group. Table 1 summarizes the results of our statistical analysis for each of our metrics.

3.2.2 Page Info interactions

Users in the control group, who saw EV UI, were significantly more likely to open the Page Info bubble. However, users in the experimental group, who did not see EV UI, were more likely to take an action in the Page Info bubble after opening it.

The experimental group opened the Page Info bubble on 0.02% of EV page loads, compared to 0.25% in the control group. Additionally, participants in our experimental group were much more likely to take an action in the Page Info bubble after opening it, across all Page Info action types.

To investigate further, we performed an additional analysis where we normalized the number of Page Info actions by the total number of EV page loads, to see if the overall number of Page Info actions taken went down in the experimental group. Table 2 shows the results of this analysis. While some Page Info actions were more common per page load in the control group, the effect sizes were very small. That is, hiding the EV UI did not make users substantially less likely to perform actions in the Page Info bubble.

Applying a Bonferroni correction for multiple testing with $m = 19$ (for each of the tests in Tables 1 and 2), the corrected significance level would instead be $\alpha = {}^{0.05}/_{m} = 0.002$. This implies that the significant results in Table 2 may be due to chance only.[8]

One possible explanation for this finding is that the large size of the EV indicator draws accidental clicks, leading users to open the Page Info bubble but not actually use it. Another hypothesis is that users notice and are curious about the EV indicator, even if it does not influence their security decisions (consistent with prior work that found that users noticed identity indicators but did not use them in their decision-making processes [28, 33, 40]). We cannot conclusively differentiate between these two hypotheses.

3.2.3 Per-site metrics

We analyzed three URL-keyed metrics (navigations, Site Engagement score, and change in Site Engagement score) on each of the top 20 EV sites. For 14 of the 20 most-visited EV origins, there were no differences with $p < 0.05$. The remaining 6 origins are shown in Table 3, each with one metric with $p < 0.05$. Five of these are very small to small negative effects on the number of navigations to the site, while one

[8]Using a Bonferroni correction allows us to control for Type I errors. We show results significant at both the $p < 0.05$ level and the multiple testing-corrected level, as the Bonferroni correction can be too conservative in cases of correlated tests.

	Control (σ)	Experiment (σ)	Δ	95% CI	p-value	Cohen's d
EV Navigations	6.18% (0.13%)	6.18% (0.13%)	-0.00	(-0.03, +0.03)	0.80	0.00
Time on EV pages (s)	2609.58 (47724.51)	2621.61 (47816.15)	+12.03	(–156.06, +180.12)	0.89	0.00
Page Ended With Tab Closed	31.64% (0.27%)	31.58% (0.27%)	-0.05	(-0.15, +0.03)	0.20	0.00
Page Ended With Back/Forward	3.61% (0.08%)	3.61% (0.08%)	+0.00	(-0.02, +0.01)	0.51	0.00
Page Reloaded	0.96% (0.05%)	0.96% (0.05%)	+0.00	(-0.02, +0.01)	0.51	0.00
Download started	3.77% (2.61%)	3.44% (1.30%)	-0.33	(-1.05, +0.38)	0.36	0.00
Form submitted	43.45% (0.97%)	43.49% (0.98%)	+0.04	(-0.30, +0.37)	0.83	0.00
CC filled	55.52% (0.79%)	55.91% (0.79%)	+0.39	(-1.21, +1.99)	0.63	0.00
Page Info opened	0.25% (0.04%)	0.02% (0.009%)	-0.23	(-0.24, -0.22)	0.00	0.09
Cookies dialog opened	0.54% (0.66%)	3.48% (0.17%)	+2.94	(+2.31, +3.57)	0.00	0.36
Changed permissions	1.26% (0.12%)	10.78% (0.34%)	+9.52	(+8.25, +10.78)	0.00	0.63
Certificate dialog opened	0.74% (0.11%)	5.52% (0.22%)	+4.78	(+3.97, +5.58)	0.00	0.39
Connection help opened	0.21% (0.04%)	1.51% (0.11%)	+1.29	(+0.89, +1.69)	0.00	0.26
Site settings opened	0.83% (0.08%)	5.80% (0.22%)	+4.97	(+4.17, +5.76)	0.00	0.49

Table 1: Summary of statistical tests for our EV field experiment metrics. The only differences that were significant at the $p < 0.05$ level were for Page Info behavior (highlighted).

	Control (σ)	Experiment (σ)	Δ	95% CI	p-value	Cohen's d
Cookies dialog opened	0.0019% (0.0026%)	0.0010% (0.0017%)	-0.001	(-0.0017, -0.00002)	0.01	0.004
Changed permissions	0.0031% (0.0028%)	0.0025% (0.0024%)	-0.0006	(-0.0015, +0.0003)	0.18	0.002
Certificate dialog opened	0.0022% (0.0026%)	0.0017% (0.0026%)	-0.0005	(-0.0014, +0.0004)	0.28	0.002
Connection help opened	0.0009% (0.0017%)	0.0005% (0.0014%)	-0.0004	(-0.0009, +0.0002)	0.17	0.002
Site settings opened	0.0028% (0.0028%)	0.0014% (0.0014%)	-0.0014	(-0.0023, -0.0006)	0.007	0.006

Table 2: Summary of our followup analysis of Page Info behavior, with counts of actions normalized by the number of EV page navigations instead. The highlighted rows were significant at the $p < 0.05$ level, but the effect sizes are negligible.

is a very small *positive* effect on the per-visit change in the Site Engagement score. However, if we apply a Bonferroni correction with $m = 60$ (three metrics checked across 20 origins), then we should instead consider a significance level of $\alpha = {}^{0.05}/_{m} = 0.0008$. With the correction, only one origin had a significant difference in user behavior: Origin 15 had 4.26 (95% CI: 2.20 to 6.32, $d = 0.24$) fewer navigations on average per user in the experimental condition.

We note that our navigation metric used here is not normalized due to limitations of the URL-keyed metrics dataset, so these results may be affected by natural variations in browsing volume between users.

4 EV survey experiments

In our EV field study, we failed to find evidence that the absence of the EV indicator influences most user behaviors. In this section, we examine two follow-up questions that were infeasible to answer via field experiment:

1. **Does the EV UI help users detect cross-jurisdiction collision attacks?** We were particularly interested in cross-jurisdiction collisions due to a recent high-profile proof of concept [11]. In this attack, two EV certificates are registered with the same legal entity name in different jurisdictions. We studied this question via survey because a field experiment would have required the browser to display incorrect information.
2. **How do users react to EV UI in modern browsers other than Chrome?** We focused on the Apple Safari browser because it recently made a significant change to its EV UI, removing the legal entity name and simply showing the domain in green (Figure 5). Because we did not have access to Safari field data, we instead conducted a survey experiment.

4.1 Methodology

We ran two online survey experiments, corresponding to the two research questions described above.

4.1.1 Questions

The surveys showed participants a login screen for a well-known financial webpage in their respective countries: PayPal in the U.S. and HSBC in the U.K. We asked participants three questions, displayed underneath the screenshot.

First, in a five-point Likert scale, we asked participants to rate their comfort level logging into the webpage: *Would you feel comfortable logging in on this website? Very comfortable / Somewhat comfortable / Neither comfortable nor uncomfortable / Somewhat uncomfortable / Very uncomfortable*

To avoid leading participants' responses, we intentionally left this question up to their interpretation and allowed them to elaborate. We next asked participants for open-ended de-

Origin	Metric	Control	Experiment	Δ	95% CI	p-value	Cohen's d
3	Navigations	18.55	14.78	-3.77	(-6.78, -0.76)	0.014	0.17
4	Navigations	25.80	23.08	-2.72	(-4.73, -0.72)	0.0078	0.10
10	Navigations	13.37	11.10	-2.27	(-4.33, -0.21)	0.031	0.14
14	Navigations	20.40	15.65	-4.74	(-8.48, -1.01)	0.013	0.21
15	Navigations	18.57	14.31	-4.26	(-6.32, -2.20)	0.00005	0.24
18	Δ Site Engagement	0.88	1.40	+0.52	(+0.09, +0.95)	0.017	0.13

Table 3: Summary of our (blinded) per-origin analysis from our UKM dataset. The included rows are the origin/metric pairs that were significant at the $\alpha = 0.05$ level. The highlighted row is the only significant result after applying a Bonferroni correction ($\alpha = 0.0008$). All the differences have at most a small effect size (Cohen's d).

tails about their reasoning: *Can you tell us why you feel that way? (If there's nothing to add, leave blank.)*

The final question appeared with the same login page screenshot, allowing users to click on it up to three times to mark the relevant sections: *Click the item(s) on the screen that make you feel that way.*

4.1.2 Participants

We recruited U.S. participants through Mechanical Turk and U.K. participants through Clickworker. We selected the U.S. and U.K. because EV usage was common in these countries (based on our dataset from Section 3), and we were unable to recruit enough participants in other countries where EV usage is common. Participants received a $.40 or €.35 incentive for participation. Our cross-jurisdiction collision survey ran from January 29 to February 3, 2019, with 592 U.S. participants and 650 U.K. participants. Our Safari EV survey ran from January 29 to February 1, 2019, with 290 U.S. participants and 305 U.K. participants.

Demographics. In both surveys, U.S. participants skewed slightly older than U.K. participants, who were overrepresented in the 18-24 age range. In the cross-jurisdiction attack survey, U.S. participants skewed slightly male (55%) and U.K. participants skewed slightly female (55%). Full demographic details can be found in the Appendix.

4.1.3 Experimental conditions

Cross-jurisdiction collision survey. In this survey, we randomly assigned participants to see one of five conditions with a screenshot of the login page, each manipulating the country code displayed in the EV indicator, as shown in Figure 4. One condition omitted the country code entirely, one showed the correct country code (US or GB), and three showed incorrect country codes (MX, RU, and BR).

Safari EV UI survey. Safari changed its EV display in macOS 10.14 to no longer display the legal entity name. In this survey, we randomly assigned participants to one of two conditions. In the first, users saw the login webpage with the EV display used in macOS 10.13, and in the second condition, users saw the EV display from macOS 10.14 (Figure 5).

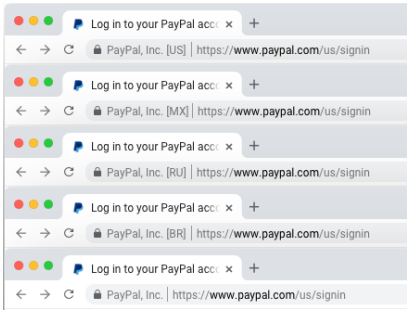


Figure 4: Five conditions shown to U.S. participants, manipulating only country code.

Figure 5: Two conditions shown to U.K. participants, manipulating display of EV to include the site's registrable domain (macOS 10.14) or EV legal entity name (as in macOS 10.13).

4.1.4 Data coding

Two researchers coded the qualitative responses on users' comfort level, with one team member (the *codemaster*) open coding the initial coding rounds, and the other iteratively providing feedback to the codemaster. In the final round of iteration, both researchers coded all responses for both surveys. Cohen's κ, a measure of inter-rater reliability, was 0.974 in the cross-jurisdiction survey (with 95.3% agreement), and 0.949 (with 97.6% agreement) in the Safari EV formatting survey, both indicating strong consistency between coders. The codemaster resolved the remaining conflicts.

4.1.5 Limitations

Artificial scenario. As with previous lab and survey studies about browser identity indicators, our surveys are an artificial scenario. This approach has limited ecological validity, as participants are not tasked with signing into a real website, nor with their real credentials, and thus they may feel less concerned than usual. However, in a more naturalistic scenario, we would expect that users would also pay less

	Cnd 1	Cnd 2	Cnd 3	Cnd 4	Cnd 5
U.S.					
Very comfortable	63%	63%	61%	56%	68%
Somewhat comfortable	30%	24%	25%	28%	21%
Neither comfortable nor uncomfortable	2%	4%	5%	3%	3%
Somewhat uncomfortable	3%	7%	6%	6%	7%
Very uncomfortable	2%	3%	3%	8%	2%
n	121	120	115	117	119
U.K.					
Very comfortable	48%	56%	46%	44%	56%
Somewhat comfortable	31%	33%	36%	39%	35%
Neither comfortable nor uncomfortable	10%	5%	3%	8%	5%
Somewhat uncomfortable	6%	4%	12%	7%	3%
Very uncomfortable	5%	2%	3%	3%	2%
n	125	132	128	132	133

Table 4: Users' comfort levels logging into a webpage with different EV country codes. Cnd 1 is the topmost variation shown in Figure 4 and Cnd 5 is the bottommost.

overall attention to security concerns because no one would ask them about their comfort level before they logged in. We therefore consider our results to describe upper bounds on how EV indicators influence user behavior.

Demographics. Since we only surveyed U.S. and U.K. participants, our results may not generalize to other contexts and cultures.

4.2 Results

Across surveys and conditions, we found that most users felt comfortable logging into each webpage, regardless of the EV UI. In nearly all cases, we found no differences among users' self-reported comfort levels with each login page.

4.2.1 Cross-jurisdiction collision survey

We found no evidence that the country code displayed in the EV indicator helps users detect a cross-jurisdiction attack.

Quantitative results. In both the U.S. and U.K., participants were most likely to say they felt "Very comfortable" logging into the webpage, regardless of the country code presented. We conducted a Kruskal-Wallis test, and in both the U.S. ($\chi^2 = 1.1783, df = 4, p = 0.8817$) and U.K. ($\chi^2 = 2.4994, df = 4, p = 0.6447$), we found no significant differences among users' comfort levels in each condition. Table 4 shows the full results.

Reasons for comfort or discomfort. When asked to identify why they felt "somewhat" or "very comfortable", participants were more likely to refer to cues in the content area, rather than Chrome UI.

Responses varied somewhat in each region. U.S. participants were most likely to describe feeling familiar with the webpage (e.g., *"PayPal is well known so it makes me feel somewhat comfortable."*), while U.K. participants most commonly pointed to an HTTPS indicator (e.g., *"the https along with the padlock in the address bar"*) but not EV-specific UI.

Participants referred to cues in the content area such as:

- familiarity with the webpage
- the page's simplicity or ease of use (e.g., *"I feel very comfortable because it is easy to understand..."*)
- the page's general design (e.g., *"A comfortable amount of white space without the page feeling empty"*)
- the page looking normal or expected (e.g., *"The sign in system here has followed a standard sign in page and gives all necessary help"*)

When referring to cues in the browser itself, participants most commonly referred to the HTTPS indicator, specifically identifying the padlock icon (e.g., *"Mainly because of the padlock on the top search bar makes me think it's secure enough to use safely"*). Participants also noted that the URL looked normal or expected (e.g., *"... the link web address doesn't look abnormal"*). They were far less likely to refer to EV UI specifically (e.g., *"The site displays that it is secure with a registered identity, PayPal Inc.... "*).

As many as 3% of U.S. participants and 14% of U.K. participants in each condition referred to the site as safe or secure, without describing their reasoning (e.g., *"It's a secure bank login page"*).

Few noticed oddities in the page's country code (no more than 8% in any U.S. condition and 5% in the U.K). Even when participants did notice, it did not necessarily make them uncomfortable (e.g., *"I never noticed the MX on a PayPal page, but it seems legit."*).

Table 5 shows a subset of results of our open-ended question about why users felt comfortable or uncomfortable.

Items on the page. When asked to "click item(s) on the page that make you feel that way", participants were most likely to click the HTTPS indicator (but not EV UI specifically), parts of the URL, or page logos. Figure 6 displays an example heatmap for these clicks. The other heatmaps can be found in the Appendix.

These results suggest that many users do use HTTPS security indicators and site URLs to determine the legitimacy of a website. However, in both qualitative and quantitative responses, almost no participants appear to notice EV UI. Additionally, these results suggest a cross-jurisdiction attack could be viable in part because users infer the legitimacy of a website from the presence of HTTPS indicators.

4.2.2 Safari EV UI survey

We found no evidence that the change in Safari's EV format affected users' comfort logging in to a webpage.

Quantitative results. In both the U.S. and U.K., in both conditions, participants were most likely to say they felt "Somewhat comfortable" or "Very comfortable" logging into the webpage. We conducted a Kruskal-Wallis test, and in both the U.S. ($\chi^2 = 0.0808, df = 1, p = 0.7762$) and U.K.

	U.S. Cnd 1	Cnd 2	Cnd 3	Cnd 4	Cnd 5	U.K. Cnd 1	Cnd 2	Cnd 3	Cnd 4	Cnd 5
n	92	120	93	93	115	83	91	81	83	74
Comfortable reasons										
I'm familiar with this website	33%	26%	31%	40%	33%	10%	7%	6%	7%	14%
I see an HTTPS indicator	32%	16%	23%	19%	17%	27%	25%	21%	23%	35%
URL looks normal	8%	8%	15%	9%	10%	1%	4%	2%	4%	4%
Page looks simple / easy to use	9%	7%	9%	10%	7%	18%	16%	9%	16%	15%
Page looks well-designed	2%	2%	0%	3%	0%	4%	8%	14%	12%	3%
I see an EV certificate	1%	1%	2%	1%	1%	1%	0%	1%	1%	1%
Uncomfortable reasons										
Country code looks strange	0%	6%	5%	8%	0%	0%	1%	5%	0%	0%
Page does not look normal	1%	1%	2%	4%	3%	1%	1%	0%	7%	3%
Page looks bland	1%	1%	4%	1%	3%	10%	2%	1%	5%	1%
URL looks odd	0%	1%	0%	1%	1%	1%	2%	2%	2%	3%
Page looks poorly-designed	0%	0%	0%	0%	0%	6%	7%	9%	7%	4%

Table 5: Sample results of the open-ended question "Can you tell us why you feel that way?" when participants were asked how comfortable they were logging in to a site. Cdn 1 is the topmost condition shown in Figure 4 and Cdn 5 is the bottommost. Full results are shown in the Appendix.

Figure 6: Example click heatmap, displaying what U.K. participants say made them feel comfortable or uncomfortable on a webpage with an RU country code in the EV indicator.

	Cnd 1	Cnd 2
U.S.		
Very comfortable	50%	47%
Somewhat comfortable	32%	30%
Neither comfortable nor uncomfortable	4%	2%
Somewhat uncomfortable	8%	16%
Very uncomfortable	6%	5%
n	142	148
U.K.		
Very comfortable	43%	42%
Somewhat comfortable	46%	39%
Neither comfortable nor uncomfortable	3%	7%
Somewhat uncomfortable	3%	11%
Very uncomfortable	4%	1%
n	152	153

Table 6: Users' comfort levels logging into a webpage with different Safari EV UIs. Cnd 1 is the variation with the site's registrable domain and Cnd 2 is the EV legal entity name.

($\chi^2 = 0.50313, df = 1, p = 0.4781$), we found no significant differences in users' comfort levels across conditions. Table 6 shows the full results.

Reasons for comfort or discomfort. Similar to the results from our cross-jurisdiction attack survey, U.S. participants were most likely to say they felt comfortable logging in because they are familiar with the webpage, while U.K. respondents were more likely to say they felt comfortable because they saw an HTTPS indicator. However, most participants in both conditions also referred to content area cues, such as the page looking as expected, or the page being simple or well-designed. Table 7 shows the full results.

Once again, as much as 6% in the U.S. and 9% in the U.K. said the website they saw is "safe" or "secure" without mentioning whether the browser or content area made them feel that way.

Participants said they felt uncomfortable logging in for several reasons, varying by region. In the U.S., participants were most likely to say they felt uncomfortable logging in because they could not see the URL (e.g., *"There's no web address present, so it could be a spoofed page"*). In the U.K. participants were most likely to say they felt uncomfortable because something in the content area was poorly-designed (e.g., *"The page looks very cold and sterile"*). Overall, however, participants were uncomfortable for very similar reasons in each region. When referring to the browser UI, they cited issues with the appearance or (in)visibility of the URL. When referring to issues with the content area, participants said the page looks bland or poorly designed.

Participants were split as to whether the EV indicator made them feel comfortable or uncomfortable, with many stating they wanted to be able to see the full URL (e.g.,

	U.S. Cnd 1	U.S. Cnd 2	U.K. Cnd 1	U.K. Cnd 2
n	115	118	95	98
Comfortable reasons				
I'm familiar with this website	40%	28%	7%	10%
I see an HTTPS indicator	25%	23%	27%	33%
Page looks simple / easy to use	8%	11%	5%	6%
Page looks normal (unclear)	7%	8%	17%	14%
It's safe / secure (unclear)	6%	2%	9%	8%
I see an EV certificate	2%	4%	2%	1%
URL looks normal	4%	0%	2%	0%
Page looks well-designed	0%	1%	12%	11%
Uncomfortable reasons				
I can't see the URL	4%	13%	3%	6%
I'm not sure if it's safe / secure (unclear)	7%	5%	3%	5%
Page looks bland	3%	7%	5%	5%
The URL looks odd	2%	1%	3%	2%
I do not see an HTTPS indicator	2%	0%	1%	0%
Page looks poorly-designed	0%	1%	9%	5%
Unclear or other	5%	12%	8%	9%

Table 7: Results of the open-ended question "Can you tell us why you feel that way?" when participants were asked how comfortable they were logging in to a site. Cdn 1 is the top condition shown in Figure 5 and Cdn 2 appears below.

"Looks like the genuine page but I'd like more reassurance of this, like being able to see the URL").

As many as 7% of U.S. participants and 5% of U.K. participants said they weren't sure if the site was safe or secure, but were unclear how (e.g., *"It doesn't look secure"*).

Items on the page. When asked to "click item(s) on the page that make you feel that way", participants were most likely to click the HTTPS indicator, as well as the page logo. Figure 7 displays a heatmap for these clicks in one condition. The other heatmaps can be found in the Appendix.

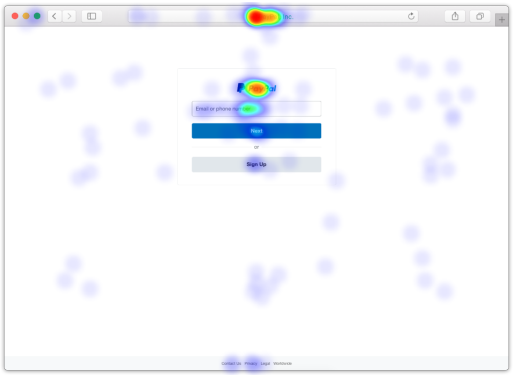


Figure 7: An example of a click heatmap from U.K. participants. This condition displayed the EV legal entity rather than a registrable domain.

5 URL highlighting survey experiment

As with the EV indicator, prior research has found that users often do not notice URLs or do not use them to make security decisions [12, 28, 40]. We conducted a survey experiment to learn whether more pronounced URL formatting changes in the browser address bar would draw attention to the URL and help users understand its security properties, but we found that these URL formatting changes were not effective.

5.1 Methodology

In this survey, we showed users a screenshot of a Google login page with a suspicious URL in the browser address bar (`accounts.google.com.amp.tinyurl.com` instead of `accounts.google.com`). We asked users to identify the website and then asked them if they would be comfortable entering their login credentials on the site.

5.1.1 Questions

The first question in the survey asked participants to identify the website in an open-ended response: *Before we move ahead, please identify the above website.* The subsequent questions asked users how comfortable they were logging in to the website and why. These questions were identical to Section 4.1.1 except that we did not ask participants to "click the item(s) on the page that make you feel that way."

5.1.2 Participants

Our survey ran from November 20 to November 21, 2018. We recruited 1,180 U.S. participants from Mechanical Turk who were paid a $.40 incentive.

Demographics. Similar to our previous U.S. surveys, the sample skewed slightly male (53%), with adults 55 and older underrepresented. Full demographic details can be found in the Appendix.

5.1.3 Experimental conditions

This survey showed participants a Google sign-in page with an incorrect URL (`accounts.google.com.amp.tinyurl.com`), simulating a phishing attack. We randomly assigned participants to one of seven conditions (Figure 8). Condition 1 (the control) used the Chrome 69 address bar UI, while other conditions attempted to draw attention to the registrable domain (`tinyurl.com`) in various ways.[9]

5.1.4 Data coding

Because there was almost no ambiguity in participant responses to our first question about the website's identity, only one researcher coded these responses. For all other questions, we coded the data as in Section 4.1.4. Based on a subsample of 100 responses coded by two security researchers,

[9] We chose these particular URL highlighting formats as we wanted to examine variants that we believed would (1) give emphasis to the registrable domain by manipulating color and spatial layout, (2) be noticeably distinguishable from the existing format but (3) not overtly distracting from browsing, so each variant could viably be deployed in the real world.

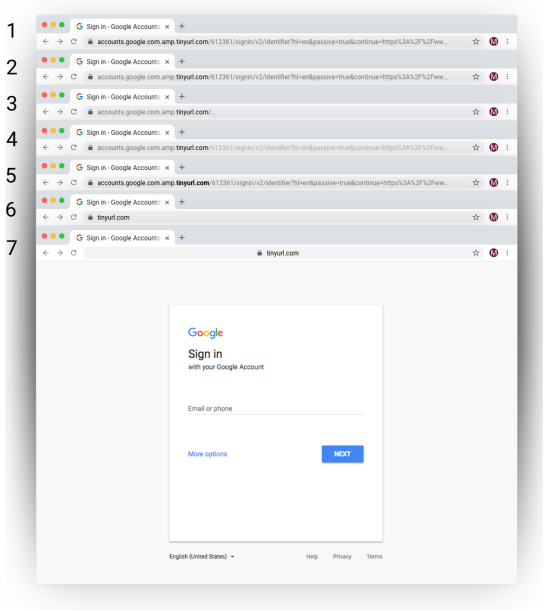


Figure 8: Conditions shown to U.S. participants, manipulating the URL display to emphasize the registrable domain.

Cohen's κ was 0.946, indicating strong agreement, with the two coders in agreement 95.4% of the time. The codemaster resolved the remaining conflicts.

5.1.5 Limitations

This survey suffers the same limitations as in Section 3.1.5: namely, an artificial scenario and limited generalizability beyond the U.S. Additionally, in this survey, participants may have responded to the novelty of the URL format, and not just the URL content, making it difficult for us to isolate the impact of the URL format alone. However, this did not appear to significantly impact our results because we did not detect any significant differences across variations.

5.2 Results

5.2.1 Website identification

Few participants noticed anything strange about the website when asked to identify it. 85% of all participants said the website was Google, when in fact, the address said `tinyurl.com`. 13% of participants correctly identified the website by its URL. 1% described both Google and TinyURL, and 1% provided a different response.

5.2.2 Comfort logging in

In all conditions, participants were most likely to say they felt comfortable logging into the webpage, despite the suspicious URL. Across the seven conditions, we found no significant differences ($\chi^2 = 2.847, df = 6, p = 0.8278$). Table 8 shows the coded results of our question about why users felt comfortable or uncomfortable logging in.

When asked why users reported feeling "somewhat" or "very comfortable", the majority of responses described looking at cues in the content area, citing that the website looked familiar (e.g., *"Because it's familiar. I've seen it plenty of times."*), or that they trust the website that appeared in the content area (e.g., *"Google is a secure company"*).

When describing discomfort, participants most commonly cited oddities with the URL (e.g., *"It seems to be an attempt to spoof Google on tinyurl"*). Relatively few participants mentioned concerns with feeling unsure how they would have navigated to this site (e.g., *"Because I have no idea how or why I'm here"*), while some described feeling unsure about the general security or safety of the site, but did not specify why (e.g., *"It's an imposter"*).

Notably, even in open-ended responses where participants appear to have been looking at the URL, they did not necessarily notice any oddities. For example, one participant reported feeling "Very comfortable" with the `tinyurl.com` URL: *"Because the URL looks like a Google page should."*

Condition 6, which showed only the registrable domain on the left of the address bar, stood out as the most distinct, with users citing oddities in the URL and generalized safety concerns at a disproportionate rate. However, the differences in comfort level between the control and this condition were not statistically significant ($\chi^2 = 0.4541, df = 1, p = 0.5004$).

6 Discussion

6.1 Summary of results

In this paper, we used large-scale field data and surveys to corroborate past results on browser identity indicators and to contribute new findings.

Our EV field experiment (Section 3) found that removing the EV UI has no effect on most user behavior metrics. However, removing EV UI did cause users to open the Page Info bubble (Figure 3) less often, and it caused a small decrease in navigations for one of the top 20 EV sites. Our experiment corroborates prior work suggesting that EV UI does not help users detect attacks [24], but at a much larger scale, with naturalistic data, and with up-to-date browser UIs. The effect on Page Info is also consistent with prior findings that users may notice EV UI but not use it in their security decisions [33].

Our EV surveys (Section 4) are the first to study cross-jurisdiction collisions and Safari's recent EV UI change. In all conditions across both surveys, EV UI did not appear to affect users' comfort levels when logging into a webpage. Our qualitative data corroborates past results that users use the content area rather than browser UI to make trust decisions [12] and that connection security indicators can be mistaken to mean that the site is safe [16]. We contribute new findings that EV indicators are likely ineffective against cross-jurisdiction collision attacks and that Safari's old and new EV UIs have similar impacts on users' comfort levels.

	Cnd 1	Cnd 2	Cnd 3	Cnd 4	Cnd 5	Cnd 6	Cnd 7
n	132	127	130	124	128	132	137
Comfortable reasons							
Looks familiar	36%	33%	35%	35%	38%	23%	32%
I trust Google	20%	17%	12%	15%	16%	16%	15%
Page looks simple / easy to use	8%	3%	8%	4%	5%	4%	4%
Site is secured or safe	5%	6%	6%	5%	6%	5%	4%
Page looks normal (unspecified)	2%	1%	0%	2%	2%	2%	1%
URL looks normal	2%	2%	0%	1%	2%	0%	0%
Uncomfortable reasons							
The URL looks funny	23%	27%	33%	27%	30%	32%	33%
I'm not sure the site is safe (unspecified)	2%	7%	2%	7%	2%	13%	4%
I'm unsure where I came from / where I am	3%	3%	2%	0%	2%	3%	1%
Unclear or other	3%	6%	3%	6%	2%	5%	9%

Table 8: Coding results of the open-ended question "Can you tell us why you feel that way?" when participants were asked how comfortable they were logging in to a site. Cdn 1 is the topmost condition shown in Figure 8 and Cdn 7 is the bottommost.

Finally, we surveyed users to determine if variations on Chrome's URL display can make it a more effective identity indicator (Section 5). None of our variations appeared to make users uncomfortable to log in to a phishing webpage. This survey corroborated prior studies showing that URLs are ineffective identity indicators [12, 28, 40], and extended them to show that several variations on browser URL display are ineffective as well. There were small but statistically insignificant differences among our variations; while a larger sample size might yield statistically significant differences, we think they are unlikely to be large effects.

6.2 Ineffectiveness of identity indicators

Removing the EV indicator did not affect most user behaviors, suggesting that an EV certificate does not provide a good defense against phishing or social engineering. While the EV UI did cause users to open Page Info more often, users did not use its functionality substantially more often. We therefore believe that users may notice the EV indicator, but do not appear to use it in making security decisions. Moreover, our survey results suggest that recent proof-of-concept attacks against EV [11] would likely be effective, and that simple UI tweaks do not make URLs an effective identity indicator either. We conclude that browser vendors should pursue more radical redesigns of their current website identity indicators if they want them to be more effective.

6.3 Guidance for designing identity indicators

Based on our experimental results and our review of prior work (Section 7), we provide the following recommendations for the design of identity indicators:

- **Prefer active, negative indicators to passive indicators.** Our UI changes failed to make the URL an effective identity indicator. Prior work has seen some success in redesigning EV indicators to make them more noticeable [33] or more understandable [8], but not better able to help users detect attacks. In contrast, active warnings like SSL errors have been successfully redesigned to reduce clickthrough rates [14]. We therefore recommend that the security community focus on triggering active warnings when a website's identity is suspicious (for example, when a domain is suspiciously similar to a popular domain), rather than relying on users to notice and act on passive identity indicators.
- **Prominent UI is an opportunity for user education.** Removing the EV indicator caused users to open the Page Info bubble less (Section 3.2.2). This effect suggests that prominent browser UI can be an opportunity to draw users' attention and educate them about the browser's identity indicators. For example, the Page Info bubble could explain the site's identity and how users should take action on it. However, we saw that in both our control and experimental groups the typical user never opened the Page Info bubble (4.65% of users in the control group opened Page Info, while 0.45% of users in the experimental group did). It is unclear if this is due to a lack of user understanding or a mismatch between users' goals and the controls provided by Page Info. Additionally, prior attempts at user education about identity indicators have been only marginally effective (e.g., [24, 28, 40]). Combined, we believe this indicates that more work is needed to understand if this approach is viable.
- **Incorporate user research in identity indicator design.** We recommend that browser vendors undergo extensive user research before launching new identity indicators, via both browser telemetry and user studies. As our work shows, both types of user research provide value: telemetry from field experiments can measure aggregate or per-site effects over large numbers of users in naturalistic settings, whereas user studies can provide insight into users' thought processes.

7 Related work

In this section, we survey related work on browser identity indicators and EV certificates.

7.1 EV effectiveness

7.1.1 User studies

Detecting fraudulent sites. In the 2000s, a number of studies analyzed how users react to EV indicators, finding that they were not effective in helping users detect phishing.

Jackson et al. [24] surveyed 27 participants about Internet Explorer 7's new EV UI. They concluded that it did not help users detect two types of phishing attacks (picture-in-picture and homograph attacks), even after receiving education about the UI.

Sobey et al. [33] analyzed Firefox 3's EV indicator as well as their own new EV design. In a lab study of 28 participants, they found that users did not notice Firefox 3's new EV indicator, but half did notice their new design. However, only a small number of participants seemed to use the newly designed indicator for decision-making.

These studies provide evidence that browser EV indicators are not effective, but they study only a small number of participants in an artificial lab scenario. Moreover, they study the very earliest EV indicators; little work has been done recently to study EV in modern browser UIs. Our work updates and expands these studies by providing large-scale in situ browser telemetry data, as well as survey data from over 1,000 participants, using modern browser UIs.

Designing EV for reassurance and understanding. Biddle et al. [8] studied Internet Explorer 7's EV indicator, comparing it to a new EV indicator of the researchers' design. Surveying 40 participants, the researchers found that their new design improved users' confidence, ease of finding information, and ease of understanding. However, they did not evaluate whether the new design helped users identify the attacks we considered. It remains an open question whether a redesigned EV indicator can effectively prevent phishing and social engineering attacks.

7.1.2 Attack proofs of concept

Researchers have recently demonstrated flaws in the EV validation procedures. The researchers obtained misleading certificates that can undermine the effectiveness of EV.

One researcher obtained a certificate for a company named "Identity Verified" [10]. This demonstrated that a malicious website could abuse the EV indicator's privileged position in browser UI to make the attack website seem more legitimate.

Another researcher obtained an EV certificate for a company named "Stripe, Inc.", mimicking the payments company but incorporated in a different state [11]. This demonstrated that EV certificates are subject to cross-jurisdiction collisions in which a user may not be able to distinguish two identical company names (one legitimate and one malicious) incorporated in different jurisdictions.

Our work is complementary to these attacks. We are primarily concerned with whether users notice and understand the EV indicator, rather than with how it can be attacked and abused. However, we do lend credence to the cross-jurisdiction collision demonstration by evaluating whether users notice cross-jurisdiction collisions (Section 4).

7.2 URL comprehension

Our work analyzes whether simple tweaks to browser URL display can help users identify fraudulent sites. Several prior studies have examined whether users understand URLs and can use them to detect attacks.

Lin et al. [28] asked 22 participants to identify fraudulent sites with and without explicit instruction to look at the browser address bar. While their user education effort was successful to an extent, it was not effective for many users and cannot be relied upon as a sole defense. Similarly, Wu et al. [39] and Dhamija et al. [12] found that neither browser address bars nor various supplemental security toolbars helped users detect phishing. In a lab study with a think-aloud protocol, Jakobsson et al. [25] concluded that users look at URLs in the process of determining whether a website is authentic, but they can be easily fooled by tricky URLs.

Xiong et al. [40] expanded Lin et al.'s work to include a control condition that did not highlight the domain in the UI, as well as a larger, more representative participant group and eye-tracking data. They found that instructing participants to look at the address bar led to a modest improvement in their ability to detect fraudulent sites, but the domain highlighting in the browser UI had no detectable effect. Their eye-tracking data suggested that explicit instructions about the browser address bar can draw users' attention to the URL, but does not give them the information or understanding that they need to draw accurate security conclusions from it.

Our work extends Xiong et al.'s study by testing multiple UI variations. Our URL formatting survey (Section 5) corroborates the existing findings that drawing users' attention to the URL bar does not help them make accurate security decisions. We contribute new findings that various UI modifications do not succeed in the goal of making the URL more noticeable and comprehensible.

7.3 Other web security UIs

Other security UIs on the web have been examined through user studies, browser telemetry, surveys, and eye-tracking.

7.3.1 Connection security indicators

Research results on browser connection security indicators have been mixed. While some studies have found that many users look at and understand them [19,37], others have found that they do not affect user behavior [12, 31]. Felt et al. [16] surveyed thousands of users to redesign connection secu-

rity indicators that met modern design constraints and better communicated the intended semantics.

Multiple studies have investigated user understanding of connection security and HTTPS, finding that users, especially those without technical backgrounds, do not have well articulated mental models for how the Internet works [26], and often conflate HTTPS and the lock icon with site security rather than connection security [38]. Krombholz et al. [27] expanded this prior work by exploring end user and administrator mental models of HTTPS, finding many misconceptions about the benefits and threat models of HTTPS among both groups. Particularly relevant for our work here, they found general distrust in HTTPS as a protocol and that security indicators are rarely part of users' mental models.

Our work contributes to this body of evidence that browser identity indicators, like connection security indicators, do not help users make security decisions. While we do not attempt to redesign identity indicators in this paper, the techniques used by Felt et al. to redesign connection security indicators could be useful for redesigning identity indicators.

7.3.2 Browser warnings and prompts

A large body of work has examined users' reactions to browser security warnings and prompts. Malkin et al. [29] and Bravo-Lillo et al. [9] conducted Mechanical Turk studies to evaluate UI changes for HTTPS warnings and plugin installation prompts, respectively. Browser security warnings have been found to have high clickthrough rates in lab studies (e.g., [12, 34, 35]), but lower in the wild [7, 14].

7.3.3 Website credibility and authenticity

Websites themselves contain security UI, including security and identity indicators. Fogg et al. [17] performed an online study to understand what makes users perceive a website as credible, and Jakobsson et al. [25] conducted a lab study to examine how users determine whether a website is authentic. These studies found that various aspects of a webpage, such as its language and spelling, can contribute to whether users perceive it as credible and/or authentic. Our survey experiments also find that users pay more attention to the website content than to browser UI when making trust decisions.

8 Conclusion

Browser identity indicators, including URLs and EV certificates, are supposed to help users identify phishing, social engineering, and other attacks, but prior lab studies and surveys suggested that older browser identity UIs are not effective security tools. In this paper, we sought to understand whether users would act on modern browser identity indicators. We provide naturalistic large-scale data about how users react to the EV indicator. We then survey thousands of users to understand the effects of recent developments in the EV ecosystem, and whether simple tweaks to browsers' URL displays can help users understand URLs better as identity indicators. We conclude that modern browser identity indicators are not effective. To design better identity indicators, we recommend that browsers consider focusing on active negative indicators, explore using prominent UI as an opportunity for user education, and incorporate user research into the design phase.

9 Acknowledgments

Thanks to Devon O'Brien, Jim Bankoski, Parisa Tabriz, Ryan Sleevi, Andrew Whalley, and Chelsea Tanaka for their support and feedback on this work.

References

[1] Domain-validated certificate. `https://en.wikipedia.org/wiki/Domain-validated_certificate`.

[2] Extended validation SSL certificate is not indicated in browser URL field. `https://github.com/brave/brave-browser/issues/3860`.

[3] Guidelines for the issuance and management of extended validation certificates. `https://cabforum.org/wp-content/uploads/CA-Browser-Forum-EV-Guidelines-v1.6.8.pdf`.

[4] Prominently show validated legal identity jurisdiction to users. `https://github.com/brave/browser-laptop/issues/791`.

[5] Public suffix list. `https://publicsuffix.org/`.

[6] URL living standard. `https://url.spec.whatwg.org/#host-registrable-domain`.

[7] AKHAWE, D., AND FELT, A. P. Alice in Warningland: A large-scale field study of browser security warning effectiveness. In *Proceedings of the 22nd USENIX Security Symposium* (2013).

[8] BIDDLE, R., VAN OORSCHOT, P. C., PATRICK, A. S., SOBEY, J., AND WHALEN, T. Browser interfaces and extended validation SSL certificates: An empirical study. In *Proceedings of the ACM Workshop on Cloud Computing Security* (2009).

[9] BRAVO-LILLO, C., KOMANDURI, S., CRANOR, L. F., REEDER, R. W., SLEEPER, M., DOWNS, J., AND SCHECHTER, S. Your attention please: Designing security-decision UIs to make genuine risks harder to ignore. In *Proceedings of the 9th Symposium on Usable Privacy and Security* (2013).

[10] BURTON, J. First part of phishing with EV. `https://www.typewritten.net/writer/ev-phishing/`.

[11] CARROLL, I. G. Extended validation is broken. `https://stripe.ian.sh/`.

[12] DHAMIJA, R., TYGAR, J. D., AND HEARST, M. Why phishing works. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2006).

[13] ESCOBAR, A. Safari Technology Preview now hides the company name (or legal entity) when showing an Extended Validation (EV) certificate, but still displays a green padlock. Progress, both in security and usability. Tweet. https://twitter.com/andrewe/status/1037737841558728706, September 2018.

[14] FELT, A. P., AINSLIE, A., REEDER, R. W., CONSOLVO, S., THYAGARAJA, S., BETTES, A., HARRIS, H., AND GRIMES, J. Improving SSL warnings: Comprehension and adherence. In *Proceedings of the 33rd Conference on Human Factors in Computing Systems* (2015).

[15] FELT, A. P., BARNES, R., KING, A., PALMER, C., BENTZEL, C., AND TABRIZ, P. Measuring HTTPS adoption on the web. In *Proceedings of the 26th USENIX Security Symposium* (2017).

[16] FELT, A. P., REEDER, R. W., AINSLIE, A., HARRIS, H., WALKER, M., THOMPSON, C., ACER, M. E., MORANT, E., AND CONSOLVO, S. Rethinking connection security indicators. In *Proceedings of the 12th Symposium on Usable Privacy and Security* (2016).

[17] FOGG, B. J., MARSHALL, J., LARAKI, O., OSIPOVICH, A., VARMA, C., FANG, N., PAUL, J., RANGNEKAR, A., SHON, J., SWANI, P., AND TREINEN, M. What makes web sites credible?: A report on a large quantitative study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2001).

[18] FRANCO, R. IE7 and High Assurance at RSA Europe. https://blogs.msdn.microsoft.com/ie/2006/10/20/ie7-and-high-assurance-at-rsa-europe/, October 2006.

[19] FRIEDMAN, B., HURLEY, D., HOWE, D. C., FELTEN, E., AND NISSENBAUM, H. Users' conceptions of web security: A comparative study. In *SIGCHI Extended Abstracts on Human Factors in Computing Systems* (2002).

[20] GOOGLE LLC. Google Chrome privacy whitepaper. https://www.google.com/chrome/privacy/whitepaper.html.

[21] HECKER, F. CAs, certificates, and the SSL/TLS UI. http://hecker.org/mozilla/ssl-ui, November 2005.

[22] HELME, S. Are EV certificates worth the paper they're written on? https://scotthelme.co.uk/are-ev-certificates-worth-the-paper-theyre-written-on/.

[23] HUNT, T. Extended Validation Certificates are Dead. https://www.troyhunt.com/extended-validation-certificates-are-dead/, September 2018.

[24] JACKSON, C., SIMON, D. R., TAN, D. S., AND BARTH, A. An evaluation of extended validation and picture-in-picture phishing attacks. In *Proceedings of the International Conference on Financial Cryptography and Data Security* (2007).

[25] JAKOBSSON, M., TSOW, A., SHAH, A., BLEVIS, E., AND LIM, Y.-K. What instills trust? A qualitative study of phishing. In *Financial Cryptography and Data Security* (2007).

[26] KANG, R., DABBISH, L., FRUCHTER, N., AND KIESLER, S. "My Data Just Goes Everywhere": User mental models of the Internet and implications for privacy and security. In *Proceedings of the 11th Symposium on Usable Privacy and Security*.

[27] KROMBHOLZ, K., BUSSE, K., PFEFFER, K., SMITH, M., AND VON ZEZSCHWITZ, E. "If HTTPS were secure, I wouldn't need 2FA": End user and administrator mental models of HTTPS. In *Proceedings of the 40th IEEE Symposium on Security & Privacy* (May 2019).

[28] LIN, E., GREENBERG, S., TROTTER, E., MA, D., AND AYCOCK, J. Does domain highlighting help people identify phishing sites? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011).

[29] MALKIN, N., MATHUR, A., HARBACH, M., AND EGELMAN, S. Personalized security messaging: Nudges for compliance with browser warnings. In *Proceedings of the 2nd European Workshop on Usable Security* (2017).

[30] PERKINS, N. Testing out #Safari in both #iOS12 and #macOSMojave and it appears that they removed the company name in the EV trust indicator and replaced it with just the URL. @iangcarroll wonder if they saw your website? Tweet. https://twitter.com/HelferNick/status/1003842702553899009, June 2018.

[31] SCHECHTER, S. E., DHAMIJA, R., OZMENT, A., AND FISCHER, I. The emperor's new security indicators. In *Proceedings of the IEEE Symposium on Security and Privacy* (2007).

[32] SIMKO, C. Why EV SSL is here to stay. https://www.globalsign.com/en/blog/why-ev-ssl-is-here-to-stay/.

[33] SOBEY, J., BIDDLE, R., VAN OORSCHOT, P. C., AND PATRICK, A. S. Exploring user reactions to new browser cues for extended validation certificates. In *Proceedings of the European Symposium on Research in Computer Security* (2008).

[34] 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 the 7th Symposium on Usable Privacy and Security* (2011).

[35] SUNSHINE, J., EGELMAN, S., ALMUHIMEDI, H., ATRI, N., AND CRANOR, L. F. Crying wolf: An empirical study of SSL warning effectiveness. In *Proceedings of the 18th USENIX Security Symposium* (2009).

[36] VERIZON. 2018 data breach investigations report. https://enterprise.verizon.com/resources/reports/dbir/.

[37] WHALEN, T., AND M. INKPEN, K. Gathering evidence: Use of visual security cues in web browsers. In *Proceedings of Graphics Interface* (2005).

[38] WU, J., AND ZAPPALA, D. When is a tree really a truck? Exploring mental models of encryption. In *Proceedings of the 14th Symposium on Usable Privacy and Security*.

[39] WU, M., MILLER, R. C., AND GARFINKEL, S. L. Do security toolbars actually prevent phishing attacks? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2006).

[40] XIONG, A., PROCTOR, R. W., YANG, W., AND LI, N. Is domain highlighting actually helpful in identifying phishing web pages? *Human Factors 59*, 4 (2017), 640–660.

Appendix

A Survey demographics

At the end of each survey we asked participants for information about their age and gender. Table 9, Table 10, and Table 11 show the demographic information for each of the three surveys.

B Full EV survey results

Figure 9 shows the full set of heatmaps for the cross-jurisdiction EV survey. Figure 10 shows the full set of heatmaps for the Safari EV survey.

Table 12 shows the full results of our open-ended coding for the cross-jurisdiction EV survey.

Gender	U.S.	U.K.
Male	55%	44%
Female	44%	55%
Other	0%	0%
Decline to answer	1%	1%
Age		
18-24	15%	31%
25-34	41%	32%
35-44	25%	20%
45-54	12%	11%
55-64	7%	4%
65+	1%	1%
Decline to answer	0%	0%
n	592	650

Table 9: Participant makeup for Chrome cross-jurisdiction EV formatting survey.

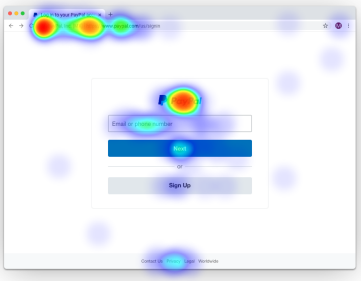
Gender	U.S.	U.K.
Male	50%	47%
Female	50%	52%
Other	0%	1%
Decline to answer	0%	0%
Age		
18-24	14%	23%
25-34	39%	32%
35-44	24%	24%
45-54	14%	11%
55-64	7%	8%
65+	2%	2%
Decline to answer	0%	0%
n	290	305

Table 10: Participant makeup for Safari EV formatting study.

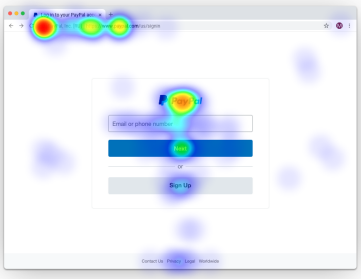
Gender	
Male	53%
Female	46%
Other	1%
Decline to answer	1%
Age	
18-24	13%
25-34	42%
35-44	24%
45-54	12%
55-64	7%
65+	2%
Decline to answer	1%
n	1180

Table 11: Participant makeup for URL formatting study.

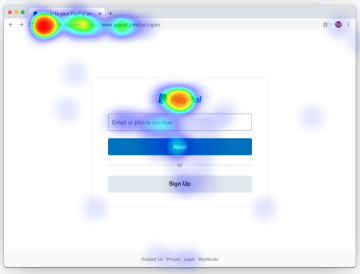
(a) US Cnd1: [US]



(b) US Cnd2: [MX]

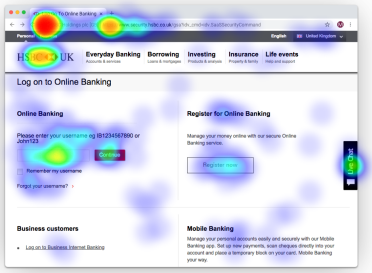


(c) US Cnd3: [RU]

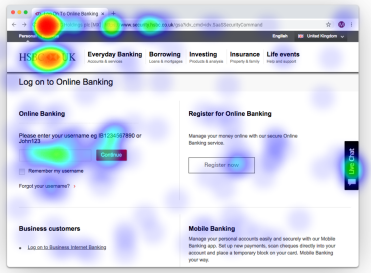


(d) US Cnd4: [BR]

(e) US Cnd5: No CC



(f) UK Cnd1: [GB]



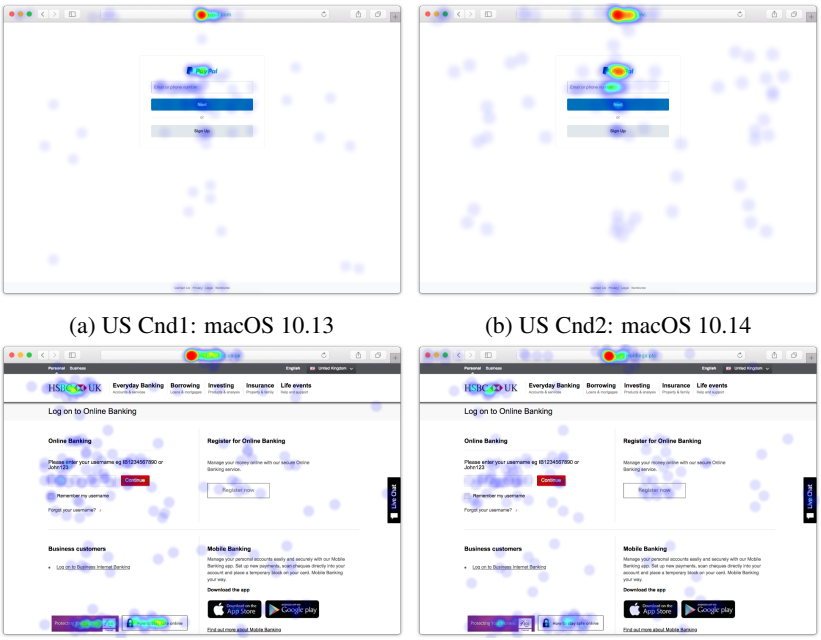
(g) UK Cnd2: [MX]

(h) UK Cnd3: [RUBR]

(i) UK Cnd4: [BR]

(j) UK Cnd5: No CC

Figure 9: Heatmaps for Chrome cross-jurisdictional EV surveys.



(a) US Cnd1: macOS 10.13

(b) US Cnd2: macOS 10.14

(c) UK Cnd1: macOS 10.13

(d) UK Cnd2: macOS 10.14

Figure 10: Heatmaps for Safari EV UI survey.

	U.S. Cnd 1	Cnd 2	Cnd 3	Cnd 4	Cnd 5	U.K. Cnd 1	Cnd 2	Cnd 3	Cnd 4	Cnd 5
n	92	120	93	93	115	83	91	81	83	74
Comfortable reasons										
I'm familiar with this website	33%	26%	31%	40%	33%	10%	7%	6%	7%	14%
I see an HTTPS indicator	32%	16%	23%	19%	17%	27%	25%	21%	23%	35%
Page looks normal (unclear)	12%	10%	10%	6%	11%	8%	10%	11%	7%	24%
URL looks normal	8%	8%	15%	9%	10%	1%	4%	2%	4%	4%
Page looks simple / easy to use	9%	7%	9%	10%	7%	18%	16%	9%	16%	15%
"It's safe / secure" (unclear)	5%	4%	8%	9%	3%	11%	16%	9%	11%	7%
Page looks well-designed	2%	2%	0%	3%	0%	4%	8%	14%	12%	3%
I see an EV certificate	1%	1%	2%	1%	1%	1%	0%	1%	1%	1%
Not asking for sensitive information	2%	3%	0%	0%	0%	5%	2%	0%	1%	0%
Uncomfortable reasons										
Country code looks strange	0%	6%	5%	8%	0%	0%	1%	5%	0%	0%
Page does not look normal	1%	1%	2%	4%	3%	1%	1%	0%	7%	3%
Page looks bland	1%	1%	4%	1%	3%	10%	2%	1%	5%	1%
Not sure if it's safe / secure (unclear)	3%	1%	1%	1%	3%	5%	1%	6%	4%	5%
Page asks for sensitive information	2%	1%	0%	3%	2%	0%	0%	0%	0%	0%
I do not see an HTTPS indicator	0%	0%	0%	3%	0%	1%	1%	0%	0%	0%
URL looks odd	0%	1%	0%	1%	1%	1%	2%	2%	2%	3%
Page looks poorly-designed	0%	0%	0%	0%	0%	6%	7%	9%	7%	4%
Lack of green security indicator	0%	0%	0%	0%	0%	0%	0%	1%	5%	0%
Unclear	3%	0%	0%	5%	0%	1%	0%	0%	4%	0%
Other	5%	1%	4%	1%	1%	2%	5%	5%	2%	1%

Table 12: Results of the open-ended question "Can you tell us why you feel that way?" when participants were asked how comfortable they were logging in to a site with different EV country code. Cdn 1 is the topmost condition shown in Figure 4 and Cdn 5 is the bottommost.