Assessing Tor's Usability as a Censorship Circumvention Tool

Linda N. Lee, David Fifield
University of California, Berkeley, {lnl, fifield}@cs.berkeley.edu

May 12, 2015

Abstract

Tor has grown beyond its original purpose of anonymity to become an important tool for Internet censorship circumvention. We describe an experiment to assess the usability of Tor Browser, a web browser with a bundled Tor proxy, for the purpose of circumvention. We focus our analysis on the browser's connection configuration dialog, which, when used correctly, enables circumvention under many of the world's most stringent censors. The experiment is aimed at evaluating whether, and how easily users can circumvent censorship using Tor Browser under several censorship environments. We present the design of the experiment, which will involve hundreds of users completing seven browsing tasks in three different adversarial settings. We describe the results of a smaller-scale pilot study, the lessons we learned from it, and how it guided the current design. The experiment we describe is scheduled to run in August 2015.

1 Introduction

We aim to improve user security by evaluating and improving the usability of security software. The topics of usability and security encompass many ideas; therefore we limit our scope and focus on a timely use case: end-user configuration of software to evade Internet censorship. In this report, we detail the design of a usability-focused user experiment, which we hope will be a benchmark for such studies and guide their development in the future.

We hold that usability is a security property. Usability considerations should be part of the design of security software. Usable software not only decreases the risk of making mistakes with potentially serious security consequences, but also provides users with a realistic alternative to substitute programs that are nice to use but less safe.

The heart of our investigation will be a large-scale user study that measures how effectively users are able to complete basic web browsing tasks in the presence of a censor. The topic of this paper is the design of the study. We will place users in a controlled network environment that simulates the kind of censorship conditions that many users around the world face daily. We will be testing specific circumvention software—namely, Tor Browser—though our larger contribution is to establish a usability testing methodology for the evaluation of this kind of software.

1.1 Censorship

Censorship and other information controls are widespread on the Internet and increasing in prevalence. Since 2011, Freedom House has published a yearly report on the state of Internet freedom in dozens of countries. The 2014 report [5] observes an overall decrease in freedom worldwide, with 36 of 65 surveyed countries decreasing in their freedom rating since 2013 (only 12 countries had an increase).

To escape restrictive networks and communicate more freely, users turn to a variety of censorship circumvention technologies. What tools are necessary depends greatly on the specifics of a particular censor; they range from simple one-hop proxies and VPNs to complicated steganographic systems. Our goal is the development of an experiment to test the usability of such tools. We focus on one particular system, Tor Browser, which bundles a variety of circumvention tools that are designed to evade different kinds of censors.

1.2 Tor

The Tor Project, best known for its anonymity network [4], has in recent years become an important player in the world of censorship circumven-

tion. In its early days, people used Tor to evade censorship, and in response censors began to block Tor itself. In response, work began on various anti-censorship technologies now known as "bridges" and "pluggable transports" [1]. As of May 2015, there are over 15,000 users accessing the Tor network at any given time using one of these anti-censorship technologies [10].

Tor Browser [9], a specially modified version of Firefox that comes with a built-in Tor proxy and pluggable transports, is the primary means through which users access the Tor network. It aims to make powerful anonymity and circumvention technology accessible even to nontechnical users. It has not undergone a formal usability assessment since its user interface was overhauled in December 2013. A central component of its user configuration interface is a panel to select bridges and pluggable transports for censorship circumvention. This selection interface is the focus of our study.

1.3 What is Required of Users

In order to circumvent censorship in their resident countries, users need to configure Tor Browser to use a bridge or pluggable transport. This requirement of the user will vary depending on their censorship environment; specifically, on whether they need to conceal the fact that they are using Tor or if Tor entry nodes have been blocked by the government. There is a configuration wizard to help guide users through setting up their connection, but they require the user to provide the IP addresses of proxies or bridges to use.

The end user generally does not care about the details of the censorship environment, and how the censorship is implemented; rather, they care about how to get to the website of their choice. In the wild, users may choose to get the required information in a variety of ways, including reading online manuals on how to configure Tor, finding out information through word of mouth, or figuring out correct configurations themselves by trial and error. We cannot meaningfully simulate the effects of these localized effects during our experiment.

However, we are testing the usability, and therefore comprehension and user-friendliness, of the Tor configuration dialog, which requires users in the most strict censorship environments to navigate through three configuration windows and provide correct answers all along the process in order to configure the connection correctly. Users must know whether their connection requires a proxy to access the Internet, the IP address of a bridge if required, and which pluggable transports are functional in their country.

2 A Pilot Tor Usability Study

We ran a pilot study to test the usability of Tor as an anonymity tool, targeting a demographic of users who may have reason to use Tor, namely, journalists who have reported in other countries. We sought to identify how well they could understand Tor's non-typical behaviors and navigate advanced settings. For the full details of this pilot experiment, see Appendix A.

During the pilot, users struggled with confusing bugs, such as a few error messages that said "Firefox" rather than "Tor Browser," disappearing buttons on the configuration window, and errors when running the program from the download archive without installing it. This resulted in our participants spending most of their time on the download and install tasks, and not what we considered the main tasks of browsing and questioning the clarity of provided security guarantees. We took the opportunity to collaborate with Tor developers to make changes in the browser and use the lessons learned to design our large-scale user study on testing Tor as a censorship circumvention tool.

2.1 Impact

Along with the Tor developers who came to observe the usability pilot, we filed a total of sixteen bug tickets to Tor. At the time of this writing, eight of the sixteen tickets have already been resolved, and more will be resolved in the future. The list of bug tickets filed can be found at `https://trac.torproject.org/projects/tor/query?keywords=~uxsprint2015`.

As of April 2015, the usability improvements resulting from our pilot study has been incorporated into the shipping release of Tor Browser. A summary of changes can be found at `https://blog.torproject.org/blog/tor-browser-45a4-released`.

Through the publicity of the usability study on the Tor blog, we were connected with a source from the Library Freedom Project [**?**], which provides train-

ing on various technologies to libraries across the country. We now collaborate with her in working on creating a user manual for Tor Browser.

2.2 Lessons Learned

During the pilot study, we were attempting to test out a comprehensive set of features for the Tor Browser. Our users had difficulties completing all the given tasks in the time allotted for the experiment. Additionally, user fatigue was very evident, with users not engaging in the later set of tasks after being tired from the download and install process. To ensure that the experiment does not cause user fatigue, we select a smaller set of features as the target for the full-scale study. After deliberating with Tor developers, and testing out various features on our own, we chose to focus on the Tor Browser connection configuration dialog because of its crucial role in censorship circumvention.

Although the data from the pilot study provided a lot of insight into what users thought processes are when completing various asks, transcribing audio recordings and syncing them to videos required much manual work. Not only did this form of data collection and processing not scale to larger numbers, we found that most users encountered similar thought processes and struggled with the same issues. Because of this, we aim to collect our future qualitative data in the form of an open-ended survey rather than through audio recordings.

3 Experiment Design

3.1 Overview

Our experiment consists of three parts: informing and motivating the participants in order to role-play in the experiment, asking our participants to complete a set of web browsing tasks which will require them to circumvent censorship using the Tor Browser, and soliciting feedback from our participants about their experience. With participants' consent, we will also record their computer screen to monitor their activity and employ a browser extension to take empirical measurements (such as idle waiting time, or how fast a particular task was completed) throughout the duration of this experiment.

Each session of 30 participants will be placed in the same simulated censorship environment. We start off the experiment by informing our participants that they are in an adversarial setting, where some websites are blocked while other websites are not. We will instruct them to visit a censored website and a non-censored website to illustrate that their devices and the network is fully functional and operating as expected, but that they are in a censorship environment where some websites will fail to load. We will take a moment to explain what Tor is, how it is generally used in this case, and ask them to complete the aforementioned set of tasks. To complete all the tasks, the participants will ultimately need to get the correct configuration settings for the country they are simulated to be in. Figure 1 shows one of the Tor configuration dialog screens that a user will need to navigate.

Then, each participant will be given about 30 minutes to complete their given Internet browsing tasks. Note that these Internet tasks themselves do not take very much time, but that the time is allotted in order for participants to complete the configuration process. We believe that this is more than enough time to test the usability of the configuration process. Even if some participants may have succeeded with more time, we consider these cases to be a failure from a usability perspective as this would require too much time from a user in the wild. Minimal interaction will be allowed with participants and the researchers. Researchers may inform participants of specific information (such as a bridge address) in a manner which does not disturb other participants (such as instructing users to visit a specific non-blocked website containing this information).

We will end the experiment with a survey of 17 questions in order to collect basic participation information, demographics of our participants, open-ended responses of their experience with censorship circumvention, and general security questions. Along with the empirical measurements, we will use this survey information to determine which configuration tasks are hard (bridges versus proxies), how many were successful, how long it took users to configure, if they understood the process, and if they were confident that it was working as expected.

3.2 Censorship Environment Simulation

We plan to simulate and test three distinct censorship environments which vary in their methods of censorship and thus require distinct responses from our participants to circumvent successfully. These

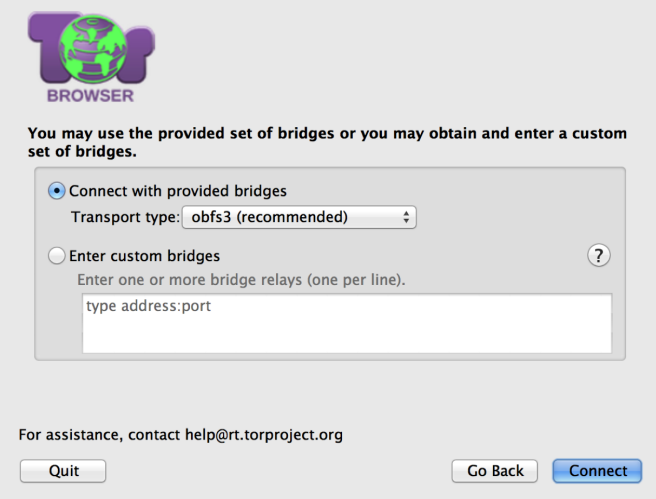

Figure 1: The fourth and final window of the Tor configuration dialog is what users will use to configure their network settings to circumvent censorship. The instructions asks users to choose a pluggable transport or enter custom bridges.

environments are designed with respect to common censorship techniques employed today, along with knowledge of how pluggable transports will need to be used. For instance, some of Tor's pluggable transports work in places that others do not. Additionally, some of them require additional information (like bridge addresses) before they work. Details of the three censorship environments are below, increasing in rough order of sophistication and difficulty to circumvent.

These environments are not meant to replicate the network environment in a particular country. Censorship environments in are complex and volatile, making it infeasible to replicating a network precisely. Moreover, precise environments is not critical to this experiment but rather environments which would require users to complete similar Tor configuration tasks before they are able to circumvent the censorship environment. We set up the following abstract environments, which are inspired by reality.

- **Corporate network.** A simulation of an enterprise or educational firewall. Blocks all services but HTTP, HTTPS, and DNS. Certain non-work-related domains, like youtube.com and torproject.org, are blocked by DNS and HTTP inspection. Blocked requests are redirected to a block page.
- **DNS-only censor.** This censor injects false replies to DNS queries for forbidden domain names (DNS poisoning), but does not do deep packet inspection on TCP streams. Unlike the corporate censor, this censor allows protocols outside a small whitelisted set.
- **Comprehensive censor.** Employs a variety of techniques, depending on the target. May do DNS poisoning, IP blocking, and inspection of TCP streams (examining the URL of HTTP requests, for example). Some domains may be blocked by DNS and deep packet inspection; others may additionally be blocked by IP address. Tor relays are blocked by IP address. Some domains, like wikipedia.org, may allow HTTP but not HTTPS. Blocked requests fail silently (no block page).

We plan to solicit feedback on the design of this experiment at the HotPETs 2015 workshop. Specifically, we would like a reliable way to target the testing of specific transports. Given that the average user does not have domain knowledge of what transports are for Tor, we wonder whether there is anything to infer from users' selection, or whether they are try transports at random?

3.3 Web Browsing Tasks

The tasks below are designed to be easy and quick to complete for an average computer user. This design is critical since the objective of the experiment is to observe how comfortable users are with circumventing censorship, which we indirectly measure by taking data regarding how participants complete the task. Giving participants tasks they are not comfortable with performing or sending them to websites which they are unfamiliar with will introduce more variables into the experiment. Specifically, we would need to separate any confusion caused by the task itself versus the task of circumventing censorship.

To create a set of tasks that an average user will be able to complete reliably, we turned to the top Alexa sites, the most popular websites on the Internet. This is an indication of representative and relevant browsing behavior. To reach the final set of website destinations, we selected sites which were popular, but also commonly censored around the world to make our study more representative. The tasks associated with each website were crafted with user familiarity in mind but narrowed down the possibilities by removing tasks which caused ethical concerns

(such as requiring a user to log in with a personal account). We believe that the tasks we will use for our study will be easy for an average user to complete.

All participants will be given the same set of tasks, regardless of their environment. Although every participant will attempt to complete the same set of tasks, the difficulty of completing these tasks will vary based on the simulated censor. Depending on the environment, some tasks may not even require the use or Tor Browser.

- Do a Google web search.
- Watch a YouTube video.
- Find the Amazon best-selling books.
- Find Yahoo's exchange rate between dollars and euros.
- Read the Wikipedia featured article.
- Find the Twitter trending topics.
- Find directions on Bing Maps.

We plant to solicit feedback on the design of this experiment at HotPETs 2015. We will solicit feedback on which additional tasks will provide additional insight, or if there are different tasks which are representative of other censorship evasion use cases.

4 Methodology

4.1 Recruitment

We plan to recruit around 200 users for the purpose of this study, making this the largest user study to date examining Tor. We do this in order to increase our chances at recruiting a diverse user population, ranging in age, gender, technical fluency, and other characteristics. Since Tor does not keep information regarding its user base, we will try to obtain a user base representative of the average Internet using-population. Additionally, recruiting a large number of participants has the benefit of ensuring enough samples in order to have significant effect sizes in each simulated censorship environment.

4.2 Quantitative and Qualitative Data

We will collect quantitative and qualitative data from our participants after they have completed their Internet tasks. This will allow us to align user activity with which simulated censorship setting, correlate performance with any demographics, get an idea of participants' motivations.

The survey consists of 17 questions total, with a basic breakdown as follows. The complete list of survey questions can be found in Appendix B.

- 2 session and participant information questions
- 4 basic demographics questions
- 5 questions regarding their experience with Tor
- 4 Internet attitude questions
- 2 technical calibration questions

We also take this time to and collect feedback on the experience as a whole and the usability of the Tor Browser. Out of the five questions, there are two multiple choice questions and three open-ended feedback questions in which our participants can write us any length of response. We opted for both concrete quantitative measurements of their experience along with free-text qualitative data to gather a broad range of information in a scalable manner.

4.3 Empirical Data

Although quantitative and qualitative data provides valuable feedback into what users perceive of their censorship circumvention experience, we will check for the validity of their statements by analyzing empirical data collected by our browser extension during the process of the experiment. Specifically, we will capture:

- A participant's computer screen.
- Configuration dialog navigation.
- Keyboard and mouse activity.
- Successfully visited websites.
- Unsuccessfully visited websites.
- Time to complete various tasks.

We hope that exploring correlations between a statement such as "I was frustrated it took so long" with task completion time, how many times a user had to repeat a dialog, or other behaviors will be an illuminating process. These correlations may hint at what makes this configuration dialog usable or not usable for users.

Additionally, users are not always reliable not complete sources of feedback. We will use the empirical data to verify the validity of their statements (for instance, if a user was reportedly frustrated but completed the task in under a minute, we may enforce our own threshold for usability) as well as performing additional analyses independent from the collected

qualitative data (such as separating time idle from time spent typing during the completion of a task).

4.4 Experimental Setup

This study will be conducted at the Experimental Social Science Laboratory (Xlab) at the University of California, Berkeley. Recruitment will be conducted through Craigslist. Although Xlab provides a portal in which researchers can recruit a pre-registered user base, we choose to not recruit through this portal, as their participants mostly consist of mostly Berkeley students. The laboratory setting consists of 36 Toshiba Tecra R850 laptops with cubicle walls separating each laptop. Since we do not have the ability to employ network-level firewalls in this laboratory setting, we will use individual host firewalls to simulate the censorship environments. We will write and employ a browser extension will record user activity, such as how many times they clicked, what websites they were successfully able to visit, and how long they took to complete each task.

The total length of the experiment, including briefing, completing the censorship circumvention tasks, exit survey, and debriefing, will take less than an hour. Participants will be compensated $30 in the form of a check or Visa debit card for their time, which is enough to cover minimum wage during their participation and their transportation costs to and from the laboratory.

5 Discussion

5.1 Limitations

One limitation of our experiment is that participants' behaviors in a laboratory setting may be different to participants' behaviors in the wild. Studies have shown that participants will act in a way which will aim to please researchers, trying to simulate what they believe are the results that we desire. Our participants may be motivated to complete the task in order to succeed at the experiment. One possible effect is that a user may have more patience to try to succeed at configuring the Tor Browser correctly during the experiment compared to if they were attempting the same task at home.

Another limitation of our experiment is that our participants might not have true interest or motivation in circumventing censorship. We could instead choose to recruit only users who have actively circumvented censorship or have an interest in Tor, but that study would also come with its own biases and limitations. We believed that both are worthwhile endeavors; however, we chose to aim for the general population. This is because we would eventually want Tor's configuration dialog to be usable for a wide array of users, including non-power users, and other types users who might not yet use Tor.

6 Future Work

With the high-level design of the experiment complete and IRB approval to conduct this experiment, the next steps are to implement the simulated censorship environments which participants will be placed in and to write the browser extension which will collect the empirical data. We wait on these tasks in order to incorporate the feedback on the design of this experiment we will receive at HotPETs 2015, which will take place during July. We plan to conduct this study starting August 2015.

We hope that this experiment will provide insight into what usability improvements can be made in order to better to make Tor Browser's advanced features more accessible to the average user. The results of this experiment can be used to inform user interaction designs which will result in shorter configuration completion time and more reliable communication of the technical task to be completed, all while the this experience does not frustrate users from future use. We believe that experiments which perform iterative A/B testing on various designs in order to increase usability to a wide, anonymous, international use base will be an interesting area of user interface design research while also making impacting changes to a widely used censorship circumvention technology.

7 Related Work

7.1 Usability for Security Software

We evaluate the usability of the configuration window's interface. Generally, usability refers to the learnability, efficiency, memorability, error rate/recoverability, and satisfaction during use. To achieve these goals, the computer human interaction community has established heuristics for user interface design: visibility of the system status, using familiar language, preventing errors, having con-

sistent visuals, flexibility and efficiency of use, error recovery, and leveraging recognition over memorization [6]. Note that these are not specific usability guidelines, but heuristics.

Usability has slightly different meanings depending on the context. Some emphasize learnability, flexibility, and ease of user. In a security context, the priorities are whatever is required for the security to be used effectively. As stated by Whitten *et al.* , a security software is usable if people who are expected to use it are reliably made aware of the security tasks to perform, are able to figure out how to successfully perform those tasks, do not make dangerous errors, and are sufficiently comfortable with the software interface [12]. Since usable security has user interface design goals than general consumer software, it likewise requires usability evaluation methods which are suitable to test these properties. Our goals are to test that the Tor Browser communicates an accurate conceptual model of the task to the users as quickly as possible, provides guidance to users at the right time, and uses warnings and other interactive tools to prevent any mistakes.

7.2 Usability and Tor

Work by Clark *et al.* to explore ways to simplify the Tor user interface has led to usability improvements [2]. Clark analyzed the challenges of using multiple Tor-related tools including TorPark, Vidalia, FoxyProxy, and TorButton to find that none are fully satisfactory from a usability perspective. Clark *et al.* also provided guidelines on how to incorporate the best aspects of each tool, while providing a set of guidelines drawn from usable security and computer human interaction. We plan to draw from Clark's guidelines as well as other works in usable security and computer human interaction to suggest any improvements to the Tor Browser.

Previous work on evaluating the usability of Tor as an anonymity system have been fruitful. Norcie *et al.* identified "stop points," or barriers to installing and using Tor to help streamline the interface. Key stop points included the ability to discriminate between anonymized and non- anonymized browser windows while using tor, unclear download and install processes, and confusion as to why the Tor behaves in peculiar ways [8]. A follow up study by Norcie *et al.* verifies that the changes recommended to the Tor Project were indeed effective at reducing stopping points mentioned in the previous work. The authors also present design heuristics for designing usable anonymous systems, such as easy installation, communicating tradeoffs made to users, and informing why, not how, precautions are taken for security guarantees [7].

A Pilot study

This pilot examined the usability of Tor as an anonymity tool with respect to journalists, a target demographic of users known to use Tor. We emphasized the targeting of a high-risk user group of Tor and to target usability changes aimed at helping those who could not afford to make mistakes. To see if those users had any potentially damaging misunderstandings or incomprehensions of Tor, we created tasks to elicit Tor's non-typical behaviors and ask our participants to navigate through more advanced settings.

A.1 Experiment Overview

Our experiment consisted of three parts. We first prepared our participants to perform a cognitive walkthrough, asked our participants to perform a list of Internet browsing tasks while performing a cognitive walkthrough, and finished with an exit survey to collect demographic information.

To first ensure that our participants are able to perform a cognitive walkthrough is, we explain what a cognitive walkthrough is and gave a demonstration of a cognitive walkthrough on a non-related dummy task (drawing a red hexagon in the computer graphics program Paint). We then asked our users to perform an unrelated Internet-based task (finding the population of Zimbabwe) using any browser of their choice while performing a cognitive walkthrough. We took this opportunity to give our participants feedback on how to properly conduct a cognitive walkthrough [11]. This process was also used to get users comfortable with talking our loud with the researcher and gauge how comfortable users were with a computer.

After the calibration task, we asked participants to complete the following Tor Browser–related tasks:

1. Download Tor (the web browser)
2. Install Tor
3. Check that Tor is working
4. Do a web search for "onions"

5. Find a YouTube video for "Ode to Joy"
6. Use "New Identity" under the onion menu
7. Explain your best understanding of all the items in the toolbar

We originally had an extra step after Step 1: "Configure Tor." We removed it after the first two participants when it became clear that the wording was confusing. The two main buttons in the configuration wizard are labeled "Connect" and "Configure"; we did not intend for users to have to go through the longer "Configure" path, but they assumed they had to because of the presence of the word "Configure" in the list of tasks.

Tasks 1–3 tested how easy it was for a user to search for, download, and install Tor Browser correctly for use and if they understood that they had done the tasks correctly. Since no software is useful if the user is not able to access it, we found it valuable to test the usability of this process. Additionally, tasks 4–6 highlighted Tor's "irregular" behavior comparatively to other web browsers. Specifically, we wanted to gauge user reactions to Tor's privacy-preserving search engine Startpage, decreased performance, and side effects of advanced settings (such as disappearing tabs upon starting a new identity). Lastly, we asked our users to interact with and describe buttons on the toolbar to gauge user comprehension of how Tor functioned.

We finished the experiment with a short exit survey which asked their basic demographics information such as their age, gender, and education. Upon finishing the survey, participants had the opportunity to ask any questions about the Tor Browser that they encountered during their walkthrough.

A.2 Methodology

We recruited five journalists from the Berkeley area by reaching out to journalist contacts. We recruited five users with an average age of 28 (σ = 13 years). Out of five participants, three were female.

We conducted the experiment on January 30-31st in Soda Hall at the University of Berkeley, California. We used a Mac laptop and a Windows laptop to give users a choice in using whatever they were comfortable with. All of our participants were Mac users. During the experiment, we recorded video data of their computer screen and audio data to capture their cognitive walkthrough. The average time to complete the given tasks, not including the cognitive walkthrough nor the exit survey, was 26 minutes ($\sigma = 7$ min).

This pilot was held in conjunction with the first Tor UX Sprint. With IRB approval, Tor developers observed the screens of our participants, which was set to broadcast to a separate room during the experiment. Audio of the cognitive walkthrough was not broadcast for privacy considerations as voice is considered an identifying characteristic.

A.3 Results

This experiment uncovered a powerful vector for attack, an unlikely source of usability issues, and how users communicated with the current interface design.

A.3.1 How to Attack Users

During the task completion process, we observed that all our participants are likely susceptible to a search engine optimization (SEO) poisoning attack during the Tor download process. Most of our participants were on autopilot and were also not aware of the possibility that their download may come from an illegitimate source. Usually, our participants picked whichever was the first search result and believed it to be a good enough source for the download:

> *"I'm going to use Chrome, because that's what I use at home... I'm just going to open it up, and I'm going to type in 'download tor' because that's what I'm supposed to do. And I'm going to click on the first one, because, at the top, which is 'Download Tor'."*—P1

> *"Okay. So I just would type in 'tor' in the uh search engine and 'Tor Project'... 'Anonymity Online'... I guess this is it. And I click. And the there is a 'Download Tor', so I am going to click this."* —P2

There was one participant who did not pick the first search result, but to save themselves time while completing the task:

> *"I am going to go to torproject.org... I am seeing the second result is 'Download Tor.' That might save a few steps, I will click on that."* —P5

Luckily, all of our users managed to download Tor browser from the legitimate source. However, defenses against SEO optimization would be prudent to employ.

A.3.2 Fighting Mac's Security Features

The biggest usability obstacle was the result of tensions between Mac's Gatekeeper code-signing feature on OS X, which prevents users from opening unsigned code downloaded from the Internet to execute on their machine unless overridden through managing system preferences. Participants could not download Tor without correctly configuring these system preferences. We found that encountering the Gatekeeper will cause frustration to most users, and may be enough to stop our participants in the download process:

> *"'The app has been modified...', blah blah blah... Oh, this is annoying. I would maybe normally stop now."* —P1

> *"'"TorBrowser" can't be opened because it's from an unidentified developer.' Grr! I express some frustration at this fact."* —P4

When users encountered this obstacle, they often returned to the download page, looking for hints or documentation. The download page did not provide help for problems such as this that are most salient immediately after downloading. As a result, Tor has developed a step-by-step visual guide on the download page to help Mac OS X users to bypassing the Gatekeeper during their download.

A.3.3 On Visual Communication

Our participants did not read much of the text provided on the download page or elsewhere during the experiment. Critical actions were predominantly guided by visual cues:

> *"...it looks like I should click there, because it is purple. No, not because it's purple, but... [laughter] because it has that little cloud with the arrow... and you can't click anywhere else to download."* —P1

> *"There's this big purple icon on the left side of the screen that says 'Download Tor Browser', so I'm going to click on that because it seems like the most obvious thing. There are a couple of other small words on the screen that I'm not going to look at, because this seems like the most obvious thing to do."* —P3

Prudently and selectively incorporating visual elements into user interface design may be the key to communicating with users and guiding correct user behaviors.

A.3.4 User Comprehension

Most of our participants did not have the expertise to understand the advanced settings for the Tor Browser:

> *"Yeah, and like the settings of your Tor Browser, I guess."* —P1

> *"But then I have toyed around before with like preferences and looking around and usually it's out of my understanding."* —P5

There was one participant who communicated a correct and thorough understanding of the new identity feature:

> *"So yeah, there was an onion. Small green onion. And yeah, I guess, it's new identity. So I'm guessing it's the same thing as with Google, when you open a page and you don't want the history to show up or something. And then you can, it's like you're disconnected from your Gmail account and everything, you can type that as a different user."* —P3

This user explained what the feature can do for him, rather than how the feature works. For the end user, the reason why a security feature is employed is what makes a feature valuable and thus motivates them to use the feature. We believe that communicating why a feature is present may be more effective for increasing use of advanced features than communicating the technical details of how a feature works.

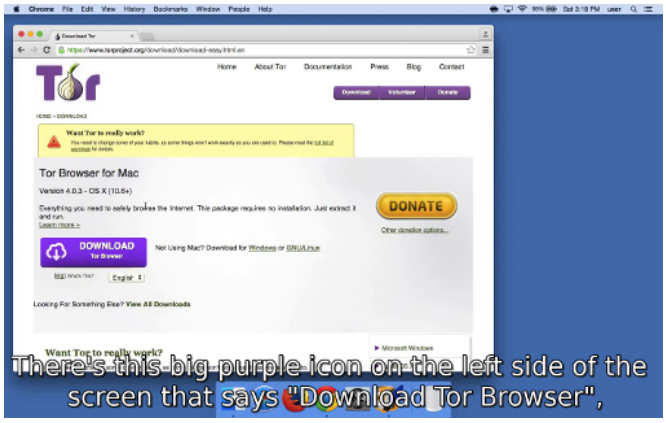

Figure 2: A participant narrates their thinking while downloading Tor Browser. *"There's this big purple icon on the left side of the screen that says 'Download Tor Browser', so I'm going to click on that because it seems like the most obvious thing. There are a couple of other small words on the screen that I'm not going to look at, because this seems like the most obvious thing to do."*

A.4 Resources

Because of the limited size of the study, it's not possible to state with confidence what fraction of users will encounter major usability obstacles. However, it was effective at discovering and demonstrating issues that are likely to cause problems for many users. For specific problems encountered by an individual user, we refer you to our resources online.

Online artifacts of our completed pilot study, such as the summary, results, and resulting browser changes are below:

- `https://trac.torproject.org/projects/tor/wiki/org/meetings/2015UXsprint`
- `https://blog.torproject.org/blog/ux-sprint-2015-wrapup`

If you would like to observe individual participants' cognitive walkthroughs, we transcribed the audio files from our experiment and synced them to the screen capture we took during the duration of the experiment. Figure 2 shows a screenshot of one of the videos.

- `https://people.torproject.org/~dcf/uxsprint2015/`

A.5 Acknowledgements

We thank the Tor Project for supporting the sprint, the participants, those who helped us recruit on short notice, and those who helped us plan and set goals and everyone who attended as a developer or observer: Arlo, Arthur, Ashkan, Griffin, Isis, Krishna, Mike, and Nima. We also thank the Tor help desk, whose #tbb-helpdesk-frequent tag helped us prioritize tickets. Special thanks go to Nima for setting up the collaboration.

B Exit Survey Questions

Our participants will be taking the following exit survey, hosted through SurveyGizmo at:

- `http://www.surveygizmo.com/s3/2085559/Tor-Usability-Survey/SG_TEST_RUN`.

Recall that the survey consists of the following 17 questions, in this order:

- 2 session and participant information questions
- 4 basic demographics questions
- 5 questions regarding their experience with Tor
- 4 Internet attitude questions
- 2 technical calibration questions

Specifically, the questions are as follows:

1. What session did you attend?
2. What is your participant ID? (This can be found on the sticker on the left hand corner of desk you are currently sitting at.)
3. What is your gender?
4. What is your age?
5. Please select your highest level of education.
6. What is your current occupation?*
7. How familiar were you with Tor prior to this study?
8. Were you confident that Tor was able to provide you with security and anonymity while completing your tasks?
9. Did anything unexpected happen while using Tor?*
10. What, if any, of the following could use improvement?*
11. Based on your experience, would you use Tor again?*
12. How many hours a week would you say you spend on the internet?

13. How concerned are you about your privacy on the Internet?
14. How concerned are you about staying anonymous on the internet?
15. How concerned are you about being censored while you browse the Internet?
16. How familiar are you with computer security?
17. What kinds of computer security software do you use?

Questions marked with asterisks are open-ended questions. All other questions are multiple choice, whether they be radio buttons, check boxes, or a drop down menu. For additional details regarding answer inputs, we refer you to our online survey linked above.

References

[1] J. Appelbaum and N. Mathewson. Pluggable transport specification, Oct. 2010. `https://gitweb.torproject.org/torspec.git/tree/pt-spec.txt`.

[2] J. Clark, P. C. Van Oorschot, and C. Adams. Usability of anonymous web browsing: an examination of tor interfaces and deployability. In *Proceedings of the 3rd symposium on Usable privacy and security*, pages 41–51. ACM, 2007.

[3] R. Dingledine and N. Mathewson. Anonymity loves company: Usability and the network effect. In *WEIS*, 2006.

[4] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. In *Proceedings of the 13th USENIX Security Symposium*, Aug. 2004. `https://svn.torproject.org/svn/projects/design-paper/tor-design.pdf`.

[5] Freedom House. Freedom on the Net. Technical report, 2014. `https://freedomhouse.org/report/freedom-net/freedom-net-2014`.

[6] J. Nielsen. 10 usability heuristics for user interface design. *Fremont: Nielsen Norman Group.[Consult. 20 maio 2014]. Disponível na Internet*, 1995.

[7] G. Norcie, J. Blythe, K. Caine, and L. J. Camp. Why Johnny can't blow the whistle: Identifying and reducing usability issues in anonymity systems. 2014.

[8] G. Norcie, K. Caine, and L. J. Camp. Eliminating stop-points in the installation and use of anonymity systems: A usability evaluation of the Tor Browser Bundle. In *5th Workshop on Hot Topics in Privacy Enhancing Technologies (HotPETS)*. Citeseer, 2012.

[9] M. Perry, E. Clark, and S. Murdoch. The design and implementation of the Tor Browser. Technical report, The Tor Project, Mar. 2013. `https://www.torproject.org/projects/torbrowser/design/`.

[10] The Tor Project. Bridge users by country, May 2015. `https://metrics.torproject.org/userstats-bridge-country.html?start=2015-02-10&end=2015-05-11`.

[11] C. Wharton, J. Rieman, C. Lewis, and P. Polson. The cognitive walkthrough method: A practitioner's guide. In *Usability inspection methods*, pages 105–140. John Wiley & Sons, Inc., 1994.

[12] A. Whitten and J. D. Tygar. Why johnny can't encrypt: A usability evaluation of pgp 5.0. In *Usenix Security*, volume 1999, 1999.