Linda Lee*, David Fifield, Nathan Malkin, Ganesh Iyer, Serge Egelman, and David Wagner

Tor's Usability for Censorship Circumvention

Abstract: Tor has grown beyond its original purpose as an anonymity tool and has become a widely-used censorship circumvention tool. This is the first study to examine Tor's usability in this role. We evaluate, design, and test the Tor configuration interface by placing users in simulated censorship environments, instructing them to use Tor to circumvent censorship, and measuring their interactions with the interface. A 16-participant qualitative user study identifies common user struggles while circumventing censorship. We use the results as feedback to redesign the configuration interface. A 114-participant quantitative user study tests the impact of our changes. We find that our changes result in a significant reduction in the time spent configuring a connections. We conclude with recommendations for changes to the current interface as well as alternative approaches to bootstrapping a connection to Tor.

Keywords: User Studies, Tor, Security, Censorship, Anonymity

DOI Editor to enter DOI
Received ..; revised ..; accepted ...

1 Introduction

Tor is an anonymity network that routes traffic through a series of relays that make it difficult to observe the source and destination [4]. Tor's anonymizing functionality also circumvents censorship. Using Tor for censorship circumvention became sufficiently common that many countries attempt to block Tor for this reason [17]. Today, Tor explicitly provides support for censorship circumvention through a network of unlisted relays as entry points into the Tor network, various methods of obfuscation to make connections to the Tor network less obvious, and advanced techniques to resist blocking.

This is the first user study that investigates the usability of Tor as a censorship circumvention tool. Our experiments are both a case study observing users circumvent censorship and a step toward helping current Tor users by making censorship circumvention easier. All Tor users benefit from improving the configuration interface, and thereby increasing the adoption of Tor as a censorship circumvention tool. Users who successfully use Tor to circumvent censorship are provided with extra security features that other censorship circumvention tools do not provide, while users who use Tor as an anonymity system benefit from an increased number of overall users on the Tor network [3].

We evaluate, redesign, and test Tor's configuration interface by measuring participants' interactions with the interface in various censorship environments. The first user study is a small-scale, qualitative experiment that collected behavioral patterns and failure cases with the interface through user observations and interviews (Section 6). This feedback was used to make changes to the interface (Section 7). The second user study is a large-scale, quantitative experiment that collected data on user interactions with the interface to quantify the impact of the design changes (Section 8).

In this paper, we contribute the following:

- 6 common challenges encountered during the configuration process and their underlying causes
- 10 changes to the configuration interface which we hope alleviates the common challenges
- 114 logs of real world user attempts to connect to Tor in three different censorship environments
- 4 reasons why users failed to connect to Tor
- 5 recommendations of configuration interface changes to help more users connect and save time
- 5 alternative approaches to bootstrapping a connection that leverage varying degrees of automation

We hope that our work helps users circumvent censorship and connect to Tor.

***Corresponding Author: Linda Lee:** University of California Berkeley, E-mail: lnl@cs.berkeley.edu
David Fifield: University of California Berkeley, E-mail: fifield@cs.berkeley.edu
Nathan Malkin: University of California Berkeley, E-mail: nmalkin@cs.berkeley.edu
Ganesh Iyer: University of California Berkeley, E-mail: ganesh.v@berkeley.edu
Serge Egelman: University of California Berkeley and International Computer Science Institute, E-mail: egelman@cs.berkeley.edu
David Wagner: University of California Berkeley, E-mail: daw@cs.berkeley.edu

2 Related Work

There have been three published user studies on Tor. Clark et al. [2] examined various deployment options for Tor Browser, such as Vidalia, Privoxy, Torbutton, and FoxyProxy, and found that none had satisfactory from a usability. Fabian et al. [6] show that Tor's added latency [5] causes users to be frustrated, cancel requests, and prevents user adoption. Norcie et al. [11] found found that 64% of users are unable to continue with installation or browsing at least once due to difficulties.

We do not know of any published usability evaluations of Tor Browser since the release of the 3.5 series in 2013, which introduced radical UI changes [12]. The most recent effort is an unpublished pilot study by Lee and Fifield [8] that tested the downloading, installation, and browsing tasks in Tor Browser. This study uncovered a number of issues [14], some of which influenced changes in Tor Browser version 4.5 and later.

Previous user studies have considered the whole browsing experience, without focusing on specific features in isolation. Our study focuses on the browser's configuration interface, which guides users through setting up components required to circumvent censorship.

3 Background

This section provides the necessary background the configuration interface relates to Tor, the network components that are involved in censorship circumvention, and the configuration settings are required to bypass levels of Tor-adverse censorship environments.

3.1 Tor, Tor Browser, and Tor Launcher

The recommended way to use Tor is through Tor Browser [13], a modified Firefox browser that includes a built-in Tor client. Tor Browser has a component called Tor Launcher that starts, stops, and otherwise controls the underlying Tor processes. Tor Launcher's graphical user interface gives access to sophisticated circumvention mechanisms, giving the user the option to configure a proxy and bridge before connecting to Tor for the first time. This is the object of our study.

In principle, the process of configuring a proxy and bridge can be automated, but the interface eschews automatic configuration through network probing in favor of guided manual configuration to give users agency in configuring their connection. Automatic configuration through network probing may put some users in certain regimes at risk. A knowledgeable user can minimize their network trace and hide that they are connecting to Tor. However, this requires that the user grapple with technical concepts such as bridges, pluggable transports, and proxies. Some of these concepts are specific to Tor and not general concepts.

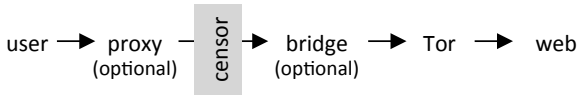


Fig. 1. The chain of components involved in connecting to a website over Tor. Most users do not need a proxy; similarly only those users who face a censor need a bridge. In the diagram, "Tor" represents all three anonymizing hops through the Tor network. We have shown the bridge as a separate component because of the special role it plays. When a bridge is used, it takes the place of the first Tor hop.

3.2 Bridges, Pluggable Transports, and Proxies

Internet censors seek to block network resources through a variety of means, such as falsifying DNS responses, blocking IP addresses, filtering keywords, and detecting protocols by deep packet inspection. Censors can block the Tor network by blocking the list of Tor relays, which are public. Blocking Tor becomes challenging when the Tor network is augmented with *bridges* and *pluggable transports*. Fig. 1 illustrates the interacting components.

Bridges are unlisted Tor relays that make it possible for a user to connect to the Tor network even if a censor blocks all publicly listed Tor relays. Pluggable transports are special protocols that run on bridges and obfuscate Tor's network protocol to make it difficult to detect. Configuring a bridge requires providing one or more "bridge lines," a specification of a bridge that includes its IP address, transport type, and other metadata. For ease of use, the interface has hard-coded options for the user to choose a group of bridges that use a particular pluggable transport. For example, choosing the hard-coded obfs3 option configures a handful of bridges that use obfs3 pluggable transport. Some censors block the IP addresses of the hard-coded bridges by looking at the source code for those addresses (except flashproxy and meek). If the built-in bridges do not work, a user can obtain bridge lines through out-of-band channels, for instance by email [1]. Fig. 2 paraphrases

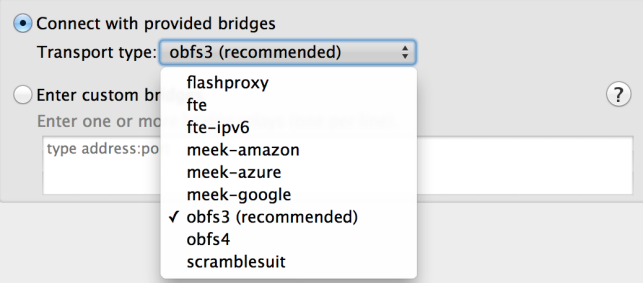


Fig. 2. Bridge selection options in Tor Launcher for Tor Browser 5.0.3. Users are not familiar with the options under "Transport type," which are various built-in circumvention technologies ("pluggable transports"). Under "Enter custom bridges," there is a space to paste in a bridge line, obtained out of band. The "Help" button displays instructions on obtaining bridge lines.

the bridge configuration options at the time of the study and lists the available pluggable transports.

In addition to Tor-specific components, a proxy may be necessary to connect outside the local network in certain managed environments, such as corporate or university networks. Configuring a proxy requires providing the proxy protocol, IP address, port, and additional optional fields. There is no assistance or automation with this process. The user must locate the proxy information and correctly input this information.

3.3 Circumventing Censorship with Tor

There are many valid configuration settings to connect to the Tor network. For instance, a user who does not need a bridge or proxy can connect to Tor with a bridge and proxy, provided that they have been configured correctly. The minimum amount of configuration required, therefore, the set of valid configuration settings, vary depending on the censorship environment. In the worst case, users must know or find out:

- Whether their Internet connection is censored
- Whether the Tor network is censored by their ISP
- Which hard-coded bridges work in their country
- If no hard-coded bridges work, how to get custom bridge information and connect to it
- Whether a proxy is required to access the Internet
- If a proxy is required, and if so, the proxy settings

Uncensored users and users in censorship environments that do not explicitly block Tor do not require a bridge or a proxy and can bypass these steps in the configuration process. Users in censorship environments that explicitly block Tor but have not blocked the hard-coded bridges in the Tor Launcher source code are required to use a hard-coded bridge or custom bridge to connect to Tor. Users in a censorship environment that has explicitly blocked Tor relays and hard-coded bridges are required to configure a meek bridge (which routes traffic through content delivery networks and therefore resists blocking through collateral damage) or configure a custom bridge to connect to Tor.

4 Evaluation Criteria and Goals

Our user studies evaluated two versions of the Tor Launcher interface. The first, which we have labeled OLD in our figures and tables, is the current version being distributed with Tor Browser. The second, NEW, is the prototype that we designed in order to provide better usability. Below are the empirical and heuristic evaluation criteria we use to assess the interfaces.

Our empirical evaluation criteria for the Tor Browser configuration interface are based on common metrics that measure ease of use:

1. **Success rate:** what percentage of users can successfully connect to Tor in a given condition. This measures how many users are able to configure their connections with the interface.
2. **Time to completion:** how long it takes between program startup and a successful Tor bootstrap, which includes time spent searching online for help and waiting for the connection to bootstrap. This measures how much time is required for a user to connect to the Tor network.
3. **Active configuration time:** how much time a user spends actively configuring network components to connect to Tor, which only includes time spent interacting with the interface, excluding the time waiting for the connection to bootstrap. This measures how much of the time spent connecting to the Tor network was spent configuring versus waiting.

An ideal interface maximizes the amount of people that successfully connect to the Tor network, while minimizing the time it takes to connect to the Tor network. A trade-off may be required between the total amount of time it takes to connect to Tor and the amount of time users interact with the interface. For instance, a process with a balanced amount of user input may be faster than a naive but completely automated process.

5 Experimental Setup

For our experiments, we use instrumented versions of Tor Browser 5.0.3, the most recent stable release at the time [7]. Though there were new releases during the experiments, we used the same version across all participants to not introduce confounding factors.

For both user studies, we simulated three censorship environments, which we refer to as E1, E2, and E3, which are in the order of increasing severity. We designed the environments to be representative of what, in our estimation, are important real-world cases of censorship. They are not intended to imitate any particular country's censorship environment.

E1: Mild censorship. (Representative of countries such as France and Australia.) The E1 environment only blocks some websites, as if by DNS poisoning. We blocked subdomains of torproject.org and wikipedia.org. (The participants' main goal in the experiment was to read a page on en.wikipedia.org.) To succeed in this environment, participants only had to click "Connect" on the first screen; additional configuration was optional.

E2: Intermediate censorship. (Representative of countries such as Tunisia.) The E2 environment blocks websites as in E1 and additionally blocks IP addresses of public Tor relays and directory authorities, simulating a censor who knows about circumvention and tries to stop it. In this environment, simply clicking "Connect" does not work; participants had to select a built-in bridge (any type other than "flashproxy" would work). A custom bridge would work but was not necessary.

E3: Comprehensive censorship. (Representative of countries such as China and Syria.) The E3 environment blocks websites and Tor relays like E1 and E2 and additionally blocks built-in default bridge IP addresses in the Tor Browser source code. To succeed in this environment, participants had to select one of the "meek" built-in bridges (no other type of hard-coded bridges would work) or acquire their own custom bridge and enter it manually.

Table 1 summarizes our simulated censorship environments. We used features of the Windows operating system to implement the above blocking behaviors. To simulate website blocking, we added entries to the hosts file, mapping domain names to the address 127.0.0.1. We used Windows Firewall rules for IP address blocking.

	E1	E2	E3
websites blocked	X	X	X
public relays blocked		X	X
default bridges blocked			X

Table 1. Summary of our simulated censorship environments. E1 only requires participants to click "Connect"; E2 requires selection of a built-in bridge; and E3 requires selection of a specific type of built-in bridge, or manual configuration of a custom bridge. E2's blocking is a superset of E1's; similarly E3's is a superset of E2's.

6 Qualitative Analysis of the Existing Interface (Study 1)

We ran a qualitative, exploratory user study to gain an understanding of how users interact with the configuration interface and what common problems may be. We did this by observing participants using the interface to circumvent censorship and interviewing them about their experience. Our observations during this study guided us in redesigning the interface and provided hypotheses to test in our quantitative study.

6.1 Inspection

We used a combination of usability inspection methods [9] to prepare for the user study. Two researchers conducted a pluralistic walkthrough and stepped through various censorship scenarios, discussing which elements would be involved for that use case, and walking through the configuration process. After compiling all the possible paths through the interface, researchers performed feature inspection, listing sequences of features used to accomplish typical tasks and taking note of long sequences or cumbersome steps. In addition, we performed a heuristic evaluation to mark likely causes of confusion for users during our study.

6.2 Recruitment

Using established practices from the field of user experience research [10], we recruited 16 participants across the three censorship environments: 5 in E1, 5 in E2, and 6 in E3. We pre-screened [15] our participants for diversity of gender, age, technical expertise, and self-reported familiarity with Tor in each simulated censorship environment for our summative usability test [16].

We recruited our users from Craigslist, vaguely asking to evaluate a piece of software. The recruitment text can be found in Appendix A. The recruitment posting contained a SurveyGizmo online survey that collected their demographics, technical expertise, and familiarity with Tor. The complete prescreening survey can be found in Appendix B. We selected participants in each censorship environment to have at least one person who has never heard of Tor, at least one person who has only heard of Tor, and exactly one person who has previously used Tor. We tried to evenly distribute any participants who had technical expertise or used particular security tools throughout the censorship environments.

Of our 16 participants, 53.3% were male. Ages ranged from 20 to 62 years ($\mu = 24.5$, $\sigma = 12.6$). 93.3% of our participants had at least a college education. 4 had used Tor before; 5 had heard of Tor but not used it; and the remaining 8 had never heard nor used it.

6.3 Procedure

The one-hour, single-participant procedure begun when a participant entered a small room with a single computer, which is equipped with Tor, Chrome, Firefox, Internet Explorer, and VLC. Each participant is informed of the risks and purpose of the study. A researcher informed the participant of the simulated censorship environment and instruct them to visit sample blocked and unblocked websites in a standard browser (Appendix C). This shows participants what a blocked site looks like in a browser. Then, the participant was asked to complete a worksheet that gives information on their censorship environment and instructed them to visit one blocked website and one non-blocked website (Appendix D). The worksheet informed the participant that the network is censored, but did not give details of what services and websites are specifically blocked. Participants were able to visit the unblocked website using any familiar browser, but had to configure Tor Browser in order to visit the blocked website.

For the blocked website, we used the main page of Wikipedia, and for the unblocked site we used the CNN homepage. We chose these two sites because of their likely familiarity to web users; our goal was to evaluate the user interface, not to test users with a browsing task. After instructions, the researchers stepped out of the room.There was no interaction between the participant and researcher for the rest of the session. Researchers watched a live video of each participant's screen from another room and saved the videos for review; the resulting videos and summaries are available from our project page. Participants had an average of 45 minutes to complete their worksheet.

After participants completed the browsing tasks or ran out of time, we interviewed them about their experience and took notes of the questions and answers. We asked questions asking about their general experience, interface features they found confusing, and feedback for improvements (Appendix E). We followed up with specific questions prompted from watching their screen. This was to verify any hypotheses we had (e.g. "they did not know what an ISP was"). We paid each participant $30 for their time.

6.4 Results

We discuss six common challenges our participants encountered during the configuration process. Participant quotations come from live transcription during the post-experiment interview and are not necessarily verbatim.

Users did not understand what proxies, bridges, and pluggable transports were. Most participants, including those pre-screened for high technical ability and previous experience with Tor, were not familiar with the vocabulary.

> P2: *"I don't know what any of those [list of bridges] means, or what that [proxy] means at all."*

> P3: *"The vocabulary is really challenging, for someone not doing IT work."*

Users did not know if they should connect directly or configure a connection. Participants incorrectly determined that a blocked torproject.org website meant that Tor relays were censored, configuring bridges and proxies when they did not need to. Other participants tried a direct connection because they did not know what to do, but configuring their connection seemed hard.

Users did not know how to choose which bridge and pluggable transport to use. On the bridge configuration screen, participants were confused by the names of the bridge transports. The most common behavior was to configure with the recommended bridge option (obfs3). If the recommended one did not work, participants did not know how to choose another.

> P8: *"I have no clue what's the difference between flash-proxy, fte, etc. I need to know why the built-in ones aren't working. And why do I need a custom bridge if there are options built in?"*

Users did not know when they are wrong. Unfortunately, many mistakes did not result in error messages, but warnings that went unnoticed. When participants did encounter an error message, they did not understand what errors meant (Fig. 3).

Users assumed they are wrong when they are right. The progress bar has a bug that causes it to update only when the level of progress increases. If progress bar reaches a 90% and fails, the next attempt have regressed to 0% and remain there until the progress surpasses 90%. Due to this, participants assumed that their subsequent attempts were wrong, even if they were right.

> P1: *"It was hard to figure out if the progress bar wasn't moving because the connection was censored, or if it was just slow."*

> P16: *"There doesn't seem to be a timeout on any of this stuff. Am I waiting long enough? It should work immediately."*

Users assume that proxy is required after a failed connection. All of our participants who failed to connect (5 of 16) failed for this reason. Many mistakenly assumed that they needed a proxy upon failure, because the interface redirects to the proxy screen (the last screen) after failure.

> P15: *"I didn't know if this computer had any proxy information. I wasn't able to find it if it did."*

From interviewing our participants, we found that these challenges are the result of these underlying causes:

Users do not know how to connect to Tor. Participants did not know the difference between a direct connection and an indirect connection to the Tor network or the difference among bridges, pluggable transports, and proxies. Participants did not know how to configure these network components without explicit additional instructions.

Feedback is too technical, missing, or misleading. Participants did not know when they failed, since certain mistakes do not trigger error warnings or timeouts (i.e. a syntactically correct but invalid proxy). If participants did see an error message, they did not understand it. The progress bar bug also caused users to wrongly assume they were not making progress since it did not give feedback on subsequent attempts.

Users do not understand censorship circumvention cues. With enough technical background, there are signals to what components are necessary and unnecessary. A connection to the Internet indicates that no proxy is required. A failed direct connection indicates a blocked Tor relay. A failed hard-coded bridge connection indicates a blocked bridge relay. However, the average user does not understand these signals.

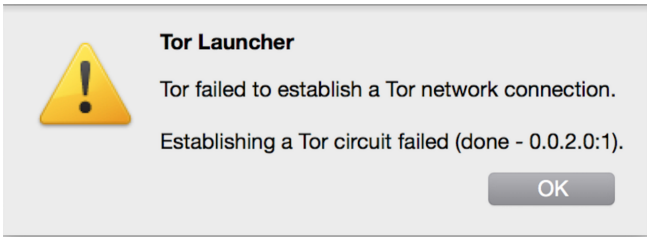


Fig. 3. An example of a technical error message which our participants did not understand.

The challenges users faced in the qualitative experiment and the respective underlying causes are used as feedback for redesigning the configuration interface.

7 Redesigning the Configuration Interface

We make ten changes to the Tor launcher configuration interface to help users connect to Tor. Since users did not know how to connect to Tor, we added advice where our users have previously struggled. Specifically, the following three changes were made to give users more information in during configuration:

1. Added instructions on what to try next on errors. When an error occurs, text advice on what to try next is shown to the user to help them recover from the error. The advice may be to try the connection again, to choose a different bridge, or to try a connection without a proxy.
2. Added guidance on choosing between connect and configure. Before, the interface only informed users that the connect option worked some of the time, but did not specify why that was so. We labeled the configure option as a manual option specifically for users in heavily censored environments.
3. Added explicit advice on choosing bridge transports. Users did not know which bridges to choose if they were in E3, which requires users to choose a meek or custom bridge. We added text that advises users to try a meek bridge if obfs3 does not work.

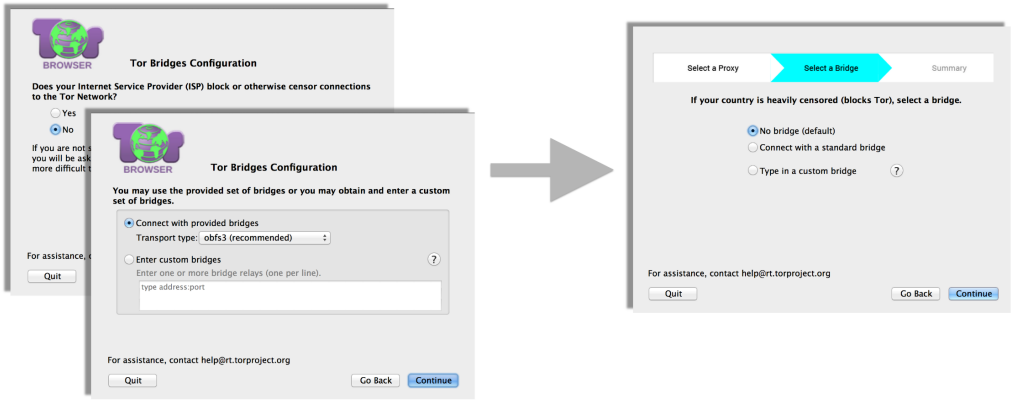


Fig. 4. In the OLD interface, users are asked, "Does your Internet Service Provide (ISP) block or otherwise censor connections to the Tor Network?" (B1). A "Yes" determines that a bridge should be configured and directs to the bridge configuration screen (B2). The NEW interface gives users advice on configuring bridges while giving the option of not configuring a bridge, on one screen (B).

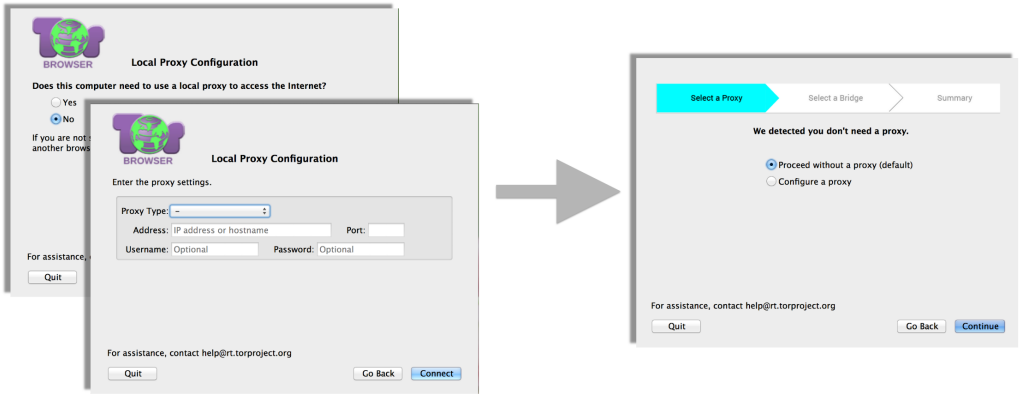


Fig. 5. In the OLD interface, users are asked, "Does this computer need a local proxy to connect to the Internet?" (P1). A "Yes" determines that a proxy should be configured and directs to the proxy configuration screen (P2). The NEW interface checks the local machine's proxy settings, and informs the user whether a proxy is required, and if so, what those settings are (P).

Users found the feedback to be more confusing than helpful, since error messages were technical and progress bar was misleading. We make the following changes to make feedback insightful and reflective of system state:

4. Progress bar feedback is accurate. We fixed the bug that caused the progress bar to not update on subsequent attempts. What users now see on the progress bar reflects the reality of the progress.
5. Made the interface text less technical. We made the text more task-centric by focusing on instructing users through the configuration process. Since users generally did not understand the technical concepts sufficiently to influence their decisions, we think giving direct guidance is a better option.
6. Added system status visibility. Before beginning any connection attempt, a summary screen displays the current bridge and proxy settings. The same information is shown on the progress screen while the connection is in progress. We switched the continuous progress bar to a discrete checkpoint-based progress indicator that shows the network components involved in connecting to the Tor network.

Our users had little technical background. We made the following changes to make the interface more compatible for the average user:

7. Eliminated technical questions. We removed the gating questions that determined whether a bridge and proxy should be configured, which were highly technical and challenging for users to answer. This are two fewer screens in the revised interface.

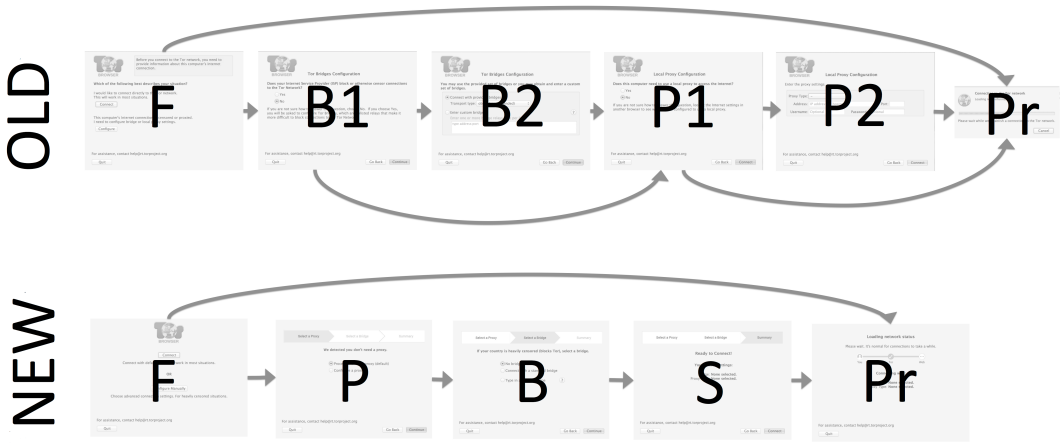


Fig. 6. Comparison of user flow in the OLD and NEW interfaces. We collapsed the two bridge screens into one and also collapsed the two proxy screens into one. We swapped the order of the bridge and proxy configuration so it matches the order of network components (compare with Fig. 1). We added a summary screen as a last step before initiating a connection.
F: first screen; **B/B1/B2**: bridge screens; **P/P1/P2**: proxy screens; **S**: summary screen; **Pr**: progress bar.

8. Added auto-detection of proxies. In principle, the interface can guess whether a proxy is needed by scanning the operating system configuration. This purely local detection does not carry any risks nor reveal to network eavesdroppers that the user is using Tor. (We simulated the auto-detection by hard-coding the fact that a proxy was not required.)
9. Switched ordering to configure proxies first. To build users' mental models, network components are now configured in a topologically sequential order, resembling Fig. 1 and the new progress screen. Previously, proxies were put after bridge configuration because only a small fraction of users require proxies. With auto-detection, configuring a proxy before a bridge burdens users less than before.

These changes result in a redesigned interface which we refer to as NEW. Note that the new interface preserves all the functionality of the old interface and still allows users to have control over their own network traffic.

8 Quantitative Analyses of the Interfaces (Study 2)

Having identified problems with the existing interface and made changes to ameliorate them, we quantify the existing problems and the impact of our changes with a study involving a larger number of users. We split participants by simulated censorship environments, and between old and new interfaces.

8.1 Setup

We ran our experiment at Xlab [18], the Experimental Social Science Laboratory at the University of California, Berkeley. Xlab has 36 Windows 7 laptops, separated by cubicle walls. Though Tor Browser runs on other operating systems, testing was only done on Windows, as a byproduct of using Xlab.

We augmented the interfaces with instrumentation to log every meaningful interaction (button presses, menu selections, screen changes). We wrote scripts to automatically set up the simulated censorship environment, install necessary software, start the video recording, and save the logs and videos. We recorded the participants' computer screens throughout the experiment to capture non-interface activity such as web searching and inspection of system networking settings.

8.2 Recruitment

We recruited 124 participants, about 20 for each condition. We recruited half of our users from Craigslist, and half of our participants from the Xlab participant pool. Although Xlab participants are not limited to UC Berkeley students and staff, a majority of the participants are from campus. For this reason, we chose to recruit half of our participants from Craigslist to ensure a diverse set of participants. The recruitment text vaguely suggests testing a piece of software, and does not require that users provide information in advance (Appendix F). Out of our 124 participants, 59 were re-

cruited from the Xlab pool and the other 65 were recruited from Craigslist. Ages ranged from 18 to 68 years ($\mu = 28.9$, $\sigma = 12$). 56.8% were male and 84.8% of our participants had at least a college education.

8.3 Procedure

The one-hour, multi-participant procedure began when all participants sat at their respective computers in Xlab. Each computer was equipped with an instrumented old or new version of Tor Browser, Chrome, Firefox, Internet Explorer, Chrome, and VLC (for screen recording). Each computer was assigned one of the six conditions in the beginning of the study. Participants were assigned to a simulated censorship environment combination at random.

A researcher informed the participants that they are in a simulated censorship environment, instructed them to visit a sample blocked website on a non-Tor browser of their choice to illustrate the situation, and asked them to complete a worksheet that asks to visit one blocked website (Appendix G). To mirror the qualitative study, we chose Wikipedia's featured article of the day as the blocked website to visit on their worksheet.

After instructions, researchers maintained minimal interactions with the participants, only answering logistical questions. The participants did not know the details of their censorship environment, only that they are being censored. Participants had 40 minutes to configure Tor Browser to circumvent the simulated censorship and visit the blocked website.

After users completed the browsing tasks, they took a short exit survey (Appendix H) that collected their demographics. All users were instructed to sit until the end of the experiment, regardless of when they had completed their task. After the 40 minutes, participants were officially informed that their time was up, and were given their payment of $30 for their time.

8.4 Results

The possible interface version and environment combinations resulted in 6 experimental conditions. We recruited 124 participants to aim for about 20 participants per condition. We filtered participants who downloaded their own version of Tor Browser or did not sign the consent form, resulting in 114 participants. Table 2 summarizes the rate of success, time to success, and active time for our participants' by condition.

	success rate after 40 minutes		median time to success	median active time
E1-NEW	**19/19**	**100%**	**0:20**	**0:06**
E1-OLD	**19/19**	**100%**	**1:01**	**0:24**
E2-NEW	**18/18**	**100%**	**3:22**	**0:40**
E2-OLD	**16/19**	**84%**	**5:00**	**2:04**
E3-NEW	**13/19**	**68%**	**20:25**	**1:56**
E3-OLD	**10/20**	**50%**	**40:08**	**9:09**

Table 2. A summary of participants' success in circumventing censorship given their simulated censorship environment and version of Tor. Those who failed to connect successfully were assigned the maximum time of 40:08.

8.4.1 Rate of Success

49 of 56 (88%) participants with the new interface successfully connected to Tor, while 45 of 58 (78%) participants with the old interface did. Due to the limited number of participants, this difference is not large enough to rule out the possibility of random chance being the cause for the difference. Appendix I details the methodology for the statistical tests used in this paper.

We added preemptive advice to the bridge configuration screen to first try an obfs3 bridge and then a meek bridge, but we suspect that most participants did not benefit from this advice since participants did not think to adjust their bridge settings upon failure. Of the 75 participants that failed to connect on the first attempt, 15 participants with the new interface and 13 participants with the old interface went back to the bridge screen and chose another hard-coded bridge. 10 of 15 in the new interface chose a meek bridge as their next bridge whereas 5 of 13 in the old interface chose meek as their next bridge, but we cannot claim choosing meek bridges is a direct result of our advice.

Table 3 shows the configuration settings of the first successful connection in each environment and interface combination.We only consider the first successful connection since many of our curious participants tried many different settings to investigate if they will work, even after they had completed the task. Recall E1 does not require users to configure a bridge, E2 requires users to configure any bridge, and E3 requires users to configure a meek or custom bridge. Note that only four of the hard-coded bridges were used to connect successfully for the first attempt were the recommended bridge and the required bridges to succeed in E3.

Our participants did not optionally configure a bridge in E1 or configure a non-recommended bridge in E2. This suggests that users will not configure optional components or deviate from the recommended settings

	E1-NEW	E1-OLD	E2-NEW	E2-OLD	E3-NEW	E3-OLD
no bridge, no proxy	17	13				
obfs3, no proxy	2	6	18	16		
meek-amazon, no proxy					7	4
meek-google, no proxy					5	4
meek-azure, no proxy					1	1
no bridge, 3rd-party proxy						1
DNF (did not finish)				3	6	10

Table 3. Bridge–proxy combinations that led to the first successful bootstrap in each environment and interface. Most E1 participants used a direct connection, but a few tried a built-in obfs3 bridge. All the E2 participants who succeeded, did so with obfs3 (the recommended bridge type)—none tried a different bridge before obfs3. All of the successful E3 participants but one used one of the meek bridges. The remaining E3 participant succeeded in an unexpected way: by searching the web for an open proxy and configuring it as the proxy setting.

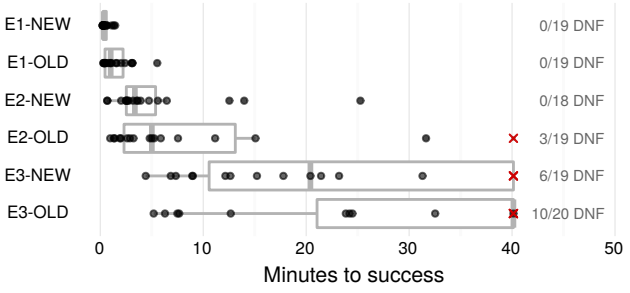


Fig. 7. Time to first success, by censorship environment and interface. The dots show the raw completion times, while the boxplots show the medians and interquartile ranges. The "DNF" figures at the right show the number of participants who did not finish in the time allotted. Here, non-finishing participants are assigned the maximum time of 40:08.

unless necessary. If the intent of the recommendation is to get as many users as possible to use the recommended bridge, this is a positive result. If the intent of the recommendation is to give pointers when users are stuck but allow the users to make their own bridge choices to diversify active transports, this is a negative result.

Only two participants chose to configure a custom bridge. Both of these participants sent an incorrectly formatted request to the automatic bridge responder, which did not reply with custom bridges as a result. Appendix J shows the malformed requests. These two participants failed to connect to the Tor network.

8.4.2 Time to Success

Time to success is defined as the time until the first successful connection to Tor. Non-finishing participants are assigned the maximum experiment time of 40:08. Our changes to the interface had a significant impact in reducing the time participants took to successfully connect to Tor (Mann–Whitney $Z = -1.84$, $p = 0.0328$, $r = 0.172$). The simulated censorship environment also had an impact; the more difficult the censorship environment, the longer participants took to configure their connection (Kruskal–Wallis $\chi^2 = 80.5$, df $= 2$, $p < 10^{-15}$). Table 2 shows the median active times and Figure 7 shows their distribution.

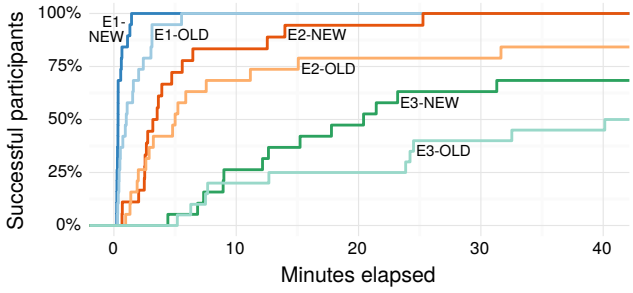
In our experiment, participants had 40 minutes to circumvent censorship. Fig. 8 shows the cumulative success rates over time. In practice, faster time to completion would mean more users will succeed, since users will give up after a while. Users in the wild will likely not be motivated to spend 40 minutes trying to configure Tor. If users were only willing to put in a minute or so of their time, users in intermediate and heavily censored environments would be unable to connect. Even if users were willing to dedicate 10 minutes to configuring their connection, most users in heavily censored environments would be unable to connect. Ideally, users should be able to connect to Tor within a few minutes, regardless of their censorship environment. We propose ideas on how to achieve this in section 8.5.3.

8.4.3 Active Time

We summarize each participant's actions throughout the experiment in Fig. 9. Each row in Fig. 9 corresponds to a participant. The bar represents a path through the interface, illustrating time spent on each screen, transitions between screens, how many attempts were made, and if they were eventually successful.

The overall time was largely dominated by the time spent waiting to connect to Tor, rather than actively configuring the interface. With correct configurations, the bootstrap process can take up to two minutes. With incorrect configurations, a lack of error messages on the progress screen caused some users to wait indefinitely. Logs show that participants spent 58% of their overall time at the progress screen. Table 4 shows the median percentage of time spent on each screen.

Perhaps a more meaningful measurement is the amount of time participants actively configured their connection (Figure 10). We define active time as the time that participants spent interacting with the in-



	1.5 m	10 m	20 m	40 m
E1-NEW	100%	100%	100%	100%
E1-OLD	58%	100%	100%	100%
E2-NEW	11%	83%	94%	100%
E2-OLD	16%	68%	79%	84%
E3-NEW	0%	26%	47%	68%
E3-OLD	0%	20%	25%	50%

Fig. 8. Cumulative success rates over time, by censorship environment and interface. We stopped participants after 40 minutes. Here, those who did not finish were assigned an arbitrarily high number greater than 40 minutes for the purposes of plotting. For example, every E1-NEW participant finished within 90 seconds, but only 58% of E1-OLD had finished by that time. Within 10 minutes, most participants in E2 had finished, along with a minority of E3 participants.

terface, excluding the time waiting for the connection to bootstrap. Active time for unsuccessful participants were calculated by subtracting the amount of time spent on the progress screen from the maximum experiment time, 40 minutes. This is an approximation, since some participants searched for help on the web after exiting the interface or on the progress screen.

We performed a one-tailed Mann–Whitney test to compare the amount of active configuration time between participants who used the new interface and the participants who used the old interface. Our changes to the interface reduced the time participants spent configuring the interface (Mann–Whitney $Z = -3.28$, $p = 0.000516$, $r = 0.307$). Table Table 2 shows the median active times and Figure 10 shows their distribution.

8.5 Discussion

We talk about the failure cases and behavioral trends observed in our quantitative user study and recommend some changes to the Tor configuration interface based on those observations.

	First	Proxy	Bridge	Progress
E1-NEW	28%	0%	0%	60%
E1-OLD	30%	0%	0%	29%
E2-NEW	6%	5%	6%	78%
E2-OLD	7%	18%	8%	45%
E3-NEW	3%	5%	5%	77%
E3-OLD	2%	12%	6%	64%

Table 4. The median percent of time spent on each screen, which is not necessarily the median absolute time spent on that screen. This percentage is computed independently for each screen; that is, a participant who spent the median percent of time on one screen may not be the same participant who spent the median percent of time on other screens. Note that the time spent on the progress bar dominates the time spent in the interface.

8.5.1 Failures

Failure is common. 17% (19 of 114) of participants were not able to successfully connect to Tor. 63% (72 of 114) of first attempts to connect failed and 79% (36 of 458) of total attempts to connect failed. Reasons for failure were determined by a combination of log processing (i.e. configuration settings on attempts) and video observation (i.e. observing what they searched for online).

Only connected directly (6/19). P73, P75, P89, P91, P106, P110 tried a direct connection. When that failed, they tried the same option, over and over, no matter how many times they failed. It was common to restart the interface, check Internet settings, and wait between subsequent attempts.

Only tried recommendations (5/19). P90, P93, P108, P111, P114, who were in E3, did not know what to do next after a direct connection and a default obfs3 bridge connection. They often tried those configurations again or gave up.

Thought that they needed a proxy (5/19). P74, P92, P105, P107, P113 assumed that they needed a proxy. They spent their time trying to configure a proxy, usually without trying other bridges.

Used the bridges auto-responder incorrectly (2/19). P94, P109, who were in E3, emailed the bridges auto-responder to get custom bridges. However, they formatted the message incorrectly and thus failed to get a response (Appendix J).

We discarded data on participants who chose to download their own version of Tor Browser online. 5 participants downloaded their own version without trying the provided one on their desktop, while others did so after feeling frustrated with the provided version. Ap-

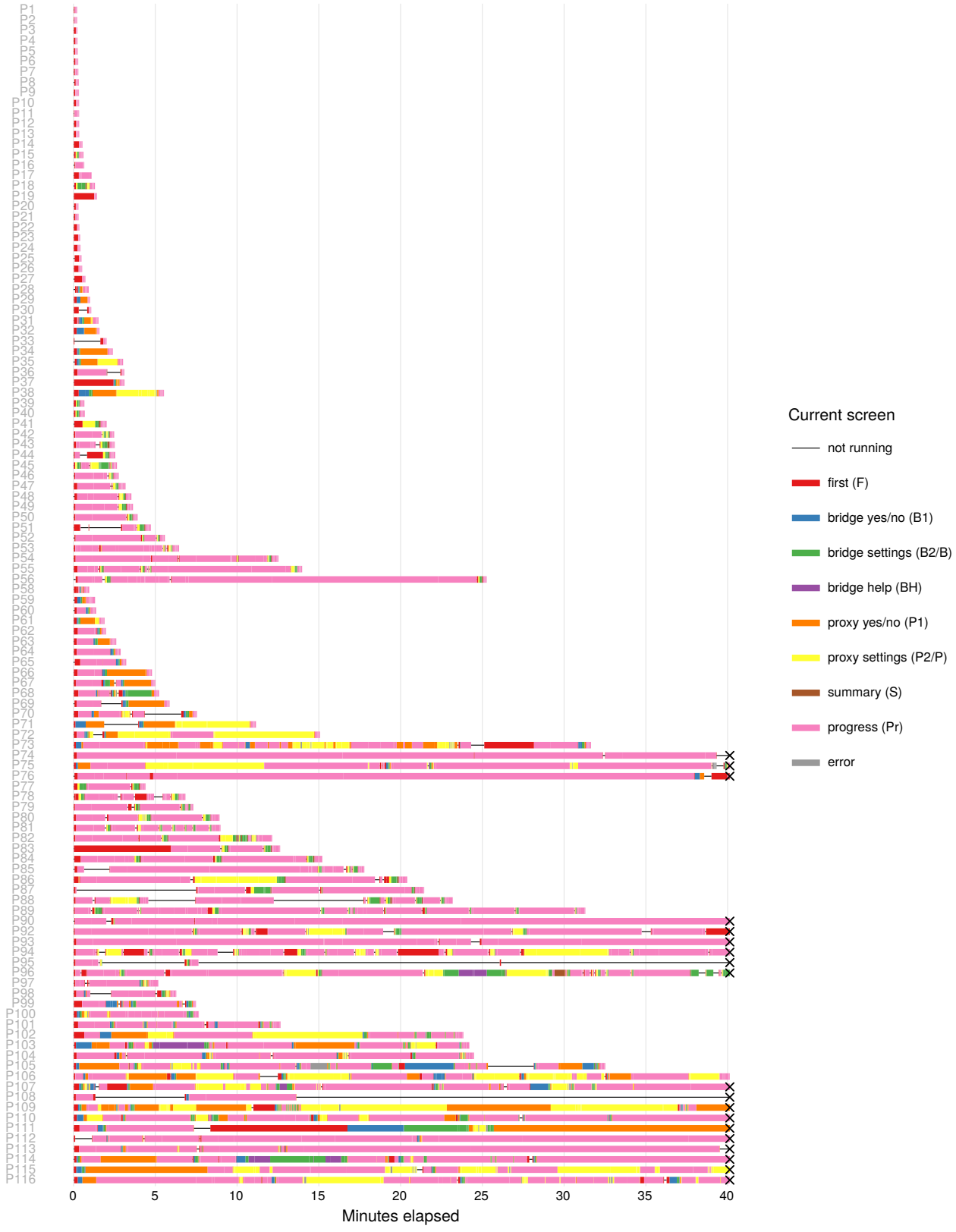


Fig. 9. Summary of participants' actions throughout the entire experiment. Different colors indicate which screen was shown at each moment. The "not running" times are those when Tor Launcher was closed; i.e., a participant was doing something else such as searching the web in another browser. The overall length of the lines show the total time to completion, except for those we cut off after approximately 40 minutes (marked with a ×).

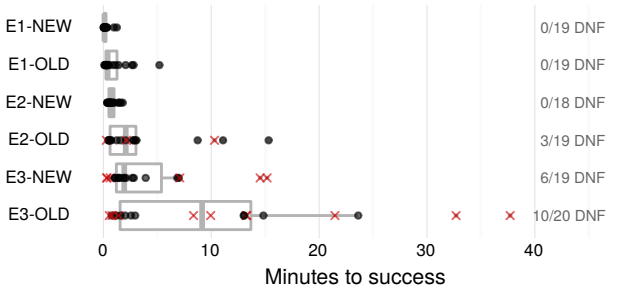


Fig. 10. Active time, by censorship environment and interface. The dots show the raw active configuration times, while the boxplots show the medians and interquartile ranges. Here, non-finishing participants' active time was computed by subtracting the amount of time spent on the progress screen from the the maximum time of 40:08.

pendix K gives the unfortunate details of where Tor was downloaded when torproject.org was blocked.

8.5.2 Observations

We noticed several behavioral patterns in our participants. These may or may not apply to the Tor user base.

Users try the easy path first. Users generally try connecting directly to the Tor network first. 81% (92 of 114) of first attempts were to connect directly, with 52% (48 of 92) direct first attempts from the new interface and 47% (44 of 93) direct first attempts from the old interface. This leaks to network eavesdroppers that they are connecting to Tor.

Users wasted time configuring proxies. This confusion prevented users from configuring their bridges correctly. Participants with the new interface (which explicitly states that they did not need a proxy) spent about the same time configuring proxies as bridges, whereas users with the old interface spent about twice as much time configuring proxies than bridges. We do not understand why users with the new interface spent time configuring proxies after explicitly being told that they do not need one.

The interface has a large influence on the bridges chosen. Participants in E3 were required to choose a particular subset of hard-coded bridges, none of which was the default bridge. After the default bridge, the most chosen bridge was the one listed first in the drop-down menu. Many bridges listed at the end of the drop-down menu were not chosen. We tried to help users by providing instructions on which bridges to choose if the default fails, but it is unclear if this advice helped. 66% (10 of 15) participants with the new interface and 38% (5 of 13) participants with the old interface chose meek as their second bridge.

People saw (almost) no error messages. Many waited for minutes at the progress screen, but never saw an error message. Error messages are intended to appear after a timeout. There were warnings, but those warnings went largely unnoticed by our participants. Across both versions of the interface, only one participant ever saw an error message. We tried to make error messages more instructive and understandable, since participants in our experiment did not encounter error messages, we cannot evaluate them.

People waited a long, long time at the progress bar. From watching the screen capture videos, we see that some participants dutifully followed the directions to wait, while others are who are uncertain wait without taking additional actions.

8.5.3 Recommendations

Based on our observations of general behavioral trends and common failure cases, we make the following recommendations to increase success rate and decrease overall time spent in the interface:

Automate configuration after failure. After one unsuccessful attempt, a user has already leaked to network eavesdroppers leaking that they are connecting to Tor. Automating the configuration process thereafter will have no increase in risk. This will specifically help users that will never try to configure their connection (the most common failure case) to succeed and drastically reduce the time spent for the 63% of participants who failed their first attempt. Users in more comprehensive censorship environments are more likely to benefit from this, as 5% of participants in the mild censorship environment, 84% participants in intermediate censorship environment, and 100% of participants in the heavy censorship environment failed the first time.

Hide infrequently used options. Telling users they did not need a proxy was not enough to deter users from configuring a proxy. Only a small fraction of the Internet population requires a proxy to connect to the Internet. Hiding the proxy screen by default and only showing it after it has been detected as

necessary might improve the user experience focusing user effort on configuring bridges.

Be explicit about recommendations. Many do not know what to do after the recommended bridge failed. The interface should mark which bridges have been tired before and recommend the user to try another bridge. (i.e. After a user tried an obfs3 bridge, the interface updates its recommendation to a meek bridge and mark the obfs3 bridge with an "X".)

Set a timeout on the progress bar. Informing users that they have failed earlier will decrease the overall time to success, create an opportunity to give suggestions, and reduce user frustration.

Have a user-tolerant bridge auto-responder. The bridge auto-responder responds to user requests for custom bridges. Users can make this request by emailing bridges@torproject.org with a non-empty title and "get bridges" in the body. Both of our users who tried to get custom bridges this way failed to format their request correctly. Having the auto-responder respond with bridges as a default or recovering from common errors (such as typing "get bridges" in the subject line) may help users succeed.

Although we cannot isolate their effects on users, we also recommend notifying a users with the configured settings before connecting and having a progress bar that illustrates when network component have successfully been configured or failed.

9 Limitations

The configuration interface uses the native operating system's elements and their respective styling, so an interface looks slightly different across different operating systems. We only tested interface on Windows machines, which were the machines available at Xlab. Participants who are not accustomed to using Windows machines may have been slower to complete the given task, but this affected all conditions equally.

We refreshed a list of Tor relay IP addresses and added them to a firewall blacklist before each session. In the first study, we neglected to block the IP addresses of the Tor directory authorities, which are the first hosts a client contacts when initiating a non-bridge connection. New relays also appeared in the network on an hourly basis, which enabled a small number of participants to succeed with a direct connection when that configuration should have failed. We believe this is acceptable because the qualitative user study was not used to quantify failures or successes, but to explore possible problems. We fixed this problem in our quantitative user study by blocking the directory authorities.

Our study was conducted in a laboratory setting, which can cause our participants to be under or over-motivated. Our participants had a monetary incentive to connect to Tor, whereas a real user in a censored environment would want to reach a particular website.

We chose to focus on more common cases, and did not simulate environments that requires users to configure a custom bridge or proxy. Although users in the wild can download interfaces that are not in their native language, we do not know how often this happens. We tested the English version of the Tor interface on English speaking participants.

10 Future Work

Our goal was to deploy impactful, tested changes the Tor configuration interface. In fact, Tor version 4.5 incorporated textual and navigational changes based on our redesigned interface. Throughout our experiments, we collaborated with Tor developers and focused on discovering changes that could be deployed right away. For this reason, we assumed that the configuration process will remain a manual process that requires user inputs, as it is currently deployed. However, we list some alternative approaches that seem worth exploring.

Automate the configuration. The most efficient way to connect as many users to the Tor network is to automatically configure their connection on start. A naive automation is to try configurations that would most likely work, in order (i.e. a direct connection, then an obfs3 connection without a proxy, then an meek bridge connection with out a proxy). This leaks to network eavesdroppers that the user is connecting to the Tor network. We do not know how much risk is associated with this approach.

Ask about the risk. An alternative to naive automation is to offer manual configuration to those that want to be more cautious and automatically configuring a connection for who are not at risk. The complication with this approach is that users may not be qualified to answer if they are at risk or may not trust Tor with this information.

Ask if users know what to do. Another alternative to naive automation is to offer manual configuration

to those that know how to configure their connection and to automate the process for users who do not know how to configure their connection. This may prevent the most mistakes, but does not account for the users' risk associated with using Tor.

Suggest configurations. A way to help users without any automation or questions is to give users information about what would work in their country. The first page of the configuration interface can show a list of countries with a corresponding recommended configuration. This approach does not require users to answer about their risk, technical ability, or location. However, it does require that the user trust the given advice and to correctly configure their settings based on this advice.

Assign configurations. This is a smart way to automate connections to the Tor network. Upon start, the interface detects proxy settings and uses them, if any. Then, all users connect to bridges that will always work (such as meek bridges), which assign them a guard relay based on their location, effectively assigning bridges for the user.

We believe that automation, asking about risk, and identifying struggling users could enable significant improvements to the configuration process.

11 Conclusion

We conducted a series of experiments to improve the Tor Browser 5.0.3 configuration interface, focusing on users who use Tor to circumvent censorship. Since connecting to the Tor network unsuccessfully indicates to network eavesdroppers that users are connecting to Tor, the configuration process is manual to allow users to have control over the network activity. Through interviews and lab studies, we find that users have difficulty configuring network components to circumvent censorship, because they do not know how censorship or Tor works. We detail the common challenges, our changes to the configuration interface address those challenges, and the recommendations we have for Tor in this paper.

Interacting with a censorship circumvention tool users can be a complex balance of leveraging user input for local information while accounting for user trust of the software and minimizing potential risks. When designing censorship circumvention tools, we caution against requiring first-time users to make decisions on software-specific notions (e.g. pluggable transports for Tor), answering questions that assume technical knowledge, or manually configuring components. We encourage the use of simulated censorship environments as a tool for user testing censorship circumvention software. Not only do simulated environments avoid rerouting traffic through a censoring country's networks, their reproducibility and stability are ideal for experiments.

12 Resources

For additional details, such as the censorship simulation code, interview transcriptions, and logs of participant interactions, we refer you to the project repository:

https://github.com/lindanlee/circumvention-ux-tor

13 Acknowledgments

I want to especially thank Rowilma del Castillo of Xlab for supporting the experiment, assisting with testing, and being flexible with non-Xlab recruiting. I have received much valuable feedback from Nima Fatemi, Isabela Bagueros, Georg Koppen, and the Tor UX team.

References

[1] BridgeDB. https://bridges.torproject.org/.

[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 R. Anderson, editor, *Proceedings of the Fifth Workshop on the Economics of Information Security*, June 2006.

[4] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. In *Proceedings of the 13th USENIX Security Symposium*, August 2004.

[5] R. Dingledine and S. J. Murdoch. Performance improvements on tor or, why tor is slow and what we're going to do about it, 2009.

[6] B. Fabian, F. Goertz, S. Kunz, S. Müller, and M. Nitzsche. Privately waiting: A usability analysis of the Tor anonymity network. In *Sustainable e-Business Management*, Lecture Notes in Business Information Processing 58, pages 63–75. Springer, 1 edition, 2010.

[7] G. Koppen. Tor Browser 5.0.3 is released, Sept. 2015. https://blog.torproject.org/blog/tor-browser-503-released.

[8] L. Lee and D. Fifield. UX Sprint 2015 wrapup. https://blog.torproject.org/blog/ux-sprint-2015-wrapup. Accessed:

2015-10-5.

[9] J. Nielsen. Usability inspection methods. In *Conference companion on Human factors in computing systems*, pages 413–414. ACM, 1994.

[10] Nielsen Norman Group. Why you only need to test with 5 users. http://www.nngroup.com/articles/how-many-test-users/.

[11] 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)*, 2012. https://petsymposium.org/2012/papers/hotpets12-1-usability.pdf.

[12] M. Perry. Tor Browser Bundle 3.5 is released, Dec. 2013. https://blog.torproject.org/blog/tor-browser-bundle-35-released.

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

[14] The Tor Project. "uxsprint2015" tickets, Mar. 2015. https://trac.torproject.org/projects/tor/query?keywords=~uxsprint2015.

[15] D. Travis. Writing the perfect participant screener. http://www.userfocus.co.uk/articles/screeners.html. Accessed: 2016-04-06.

[16] User Experience Professionals' Association. Summative usability testing. http://www.usabilitybok.org/summative-usability-testing. Accessed: 2016-04-06.

[17] P. Winter and S. Lindskog. How the Great Firewall of China is blocking Tor. *Free and Open Communications on the Internet*, 2012. https://www.usenix.org/system/files/conference/foci12/foci12-final2.pdf.

[18] Xlab: Experimental Social Science Laboratory. https://xlab.berkeley.edu/.

A Qualitative User Study Recruitment Posting

We are recruiting participants for an in-person research study at the University of California, Berkeley. You will need to come in to our lab and perform tasks on a computer for an hour or less. You will be compensated $30 for participating. No special knowledge and no technical experience is required. If you are interested, fill out the survey at *<survey link>*.

B Qualitative User Study Prescreening Survey

We are recruiting participants for an in-person research study at the University of California, Berkeley. You will need to come in to our lab and perform tasks on a computer for an hour or less. You will be compensated $30 for participating. No special knowledge and no technical experience is required.

1. Please select when you are available. We will assign you an hour experiment time slot during one of those times.
2. I am able to provide my own transportation to the University of California, Berkeley campus.
3. Thank you for your interest! Please provide an email address where we can contact you to share more logistical details.
4. we are looking for a very small number of participants, so unfortunately, we may not be able to accommodate everyone who applies. Would you like us to let you know about future opportunities?
5. What is your gender?
6. What is your age?
7. Please select your highest completed (or current) level of education.
8. What is your occupation?
9. Do you speak any languages other than English fluently?
10. If you have a personal computer, what kind do you use?
11. Which of the following terms have you heard of? *<answer choices: a checkboxlist of the the following terms: malware, proxy services, phishing, SSL, X.511 certificates, Tor>*
12. How often do you use the following software or features? *<answer choices: a grid of radio buttons. Software/features (rows): HTTPS on web pages, proxies or other censorship circumvention tools, virtual private networks (VPN), file or whole-disk encryption, anonymity systems (e.g., Tor), email encryption (e.g., PGP), chat or instant messaging encryption, voice communication encryption. Frequency (columns): never, less than once a month, a few times a month, several times a week, daily.>*

Thank you for filling out this form. You are now done!

C Qualitative User Study Introduction Script

Imagine you live in an oppressive country that censors part of the Internet. We have simulated this in the laboratory by blocking certain websites and services. The purpose of this experiment is to evaluate the use of Tor browser, which is a browser that can circumvent censorship and let you visit blocked websites. Currently, torproject is blocked (you can check this by going to torproject.org on a standard browser, like Firefox, Chrome, or Internet Explorer).

To circumvent censorship successfully, you will need to set up Tor browser correctly and use it to get to Wikipedia. If you are able to reach the website, then you know that you have successfully circumvented censorship. Fill out the question on the worksheet. This isn't intended to be hard, just write what you see. We want to just check you saw the website.

Before you start, do you have any questions about what you are asked to do?

D Participant Worksheet Text

Imagine you live in an oppressive country that censors part of the Internet. We have simulated this in the laboratory by blocking certain websites and services. The purpose of this experiment is to evaluate the use of Tor browser, which is a browser that can circumvent censorship and let you visit blocked websites. For instance, www.torproject.org is blocked. Check this by going to the site on a standard browser, like Firefox, Chrome, or Internet Explorer. It will fail to load, when you can visit other sites.

To complete this worksheet, you will need to set up Tor browser (on your desktop) correctly and use it to get to blocked site. If you can visit wikipedia, then you know that you have successfully circumvented censorship.

E Post-Experiment Standard Interview Questions

We asked our participants these questions after they were given time to configure Tor Browser.

1. Can you talk us through what you did along with what you were thinking at the time?
2. What was most challenging part of connecting?
3. Were there any unfamiliar terms?
4. How did you decide which options to choose?
5. What did you think about using Tor?
6. What is one change you would recommend?
7. Did you need any additional information?

In addition to these questions, we asked our participants about specific questions based on their observation, usually regarding a specific choice in action, a particular screen they seemed stuck on, and any errors they encountered during the configuration process.

F Quantitative User Recruitment Posting

We are recruiting up to 40 participants for a user study at UC Berkeley. The experiment will involve basic Internet browsing tasks. You are not eligible if you have participated in our previous sessions.

Payment: $30 Amazon gift card
Duration: 1 hour
Where: Xlab at Hearst Memorial Gymnasium

<list of sessions>

To be eligible, you must be an adult (18 or older). This is to comply with university policies on research.

If you are interested: 1. Email lnl@berkeley.edu with the sessions you are able to attend. We will confirm your participation and assign you a session. 2. Come to Xlab at the appointed time for the experiment.

G Quantitative User Study Introduction Script

Imagine you live in an oppressive country that censors part of the Internet. We have simulated this in the laboratory by blocking certain websites and services. The purpose of this experiment is to evaluate the use of Tor browser, which is a browser that can circumvent censorship and let you visit blocked websites. Currently, torproject is blocked (you can check this by going to tor-

project.org on a standard browser, like Firefox, Chrome, or Internet Explorer).

To circumvent censorship successfully, you will need to set up Tor browser correctly and use it to get to Wikipedia. Tor is already installed for you. On the desktop, you should see a globe icon that says "Start Tor Browser." If you are able to reach the website, then you know that you have successfully circumvented censorship. Fill out the question on the worksheet. This isn't intended to be hard, just write what you see. We want to just check you saw the website.

Afterward, we ask you to take a short survey to collect some information about you. The link is also on your worksheet. We will give you time to complete this task. If you finish early, we ask that you sit at your desk until the remainder of the hour. Since we are recording your screen, we ask that you don't do anything personal afterward, like checking your email.

Before you start, do you have any questions about what you are asked to do?

H Quantitative User Study Exit Survey

We'd like to know more about you. All of your answers will be stored separately from any identifying information in order to protect your confidentiality.

This survey is part of a research project being conducted by the University of California, Berkeley. If you have any questions about your rights or treatment as a research participant in this study, please contact the University of California at Berkeley's Committee for Protection of Human Subjects at 510-642-7461, or email subjects@berkeley.edu. If you agree to participate, please click Next below.

1. What is your participant ID? (This can be found on the sticker on the left hand corner of the desk you are currently sitting at.)
2. What is your gender?
3. What is your age?
4. Please select your highest completed (or current) level of education.
5. What is your current occupation?

Thank you for participating in our experiment. You are now done! Please sit at your desk for the remainder of the experiment. Our researchers will formally announce the end of the experiment.

I Statistical Tests

From our measurements, we observe that participants with the new interface have a higher rate of success, succeed in less time, and configure their interface in less time. We do not find that the increased rate of success with the new interface or the decreased time to success with the new interface significant. That is, random chance can account for the difference. We do find the the decreased active configuration time to be significant. We describe the methodology for the statistical tests used in this paper to determine the impact of the interface on success rate, time to success, and active time.

Each participant had a boolean variable indicating a successful connection to Tor. Rates of success were calculated by dividing the number of participants who succeeded the condition over the total number of participants in the condition. This gave us six rates of success. E1-NEW, E2-NEW, and E3-NEW rates of success were compared against E1-OLD, E2-OLD, and E3-OLD. We used a Pearson's Chi-squared test to test the significance of success rates. The difference between success rates of participants with the new interface and participants with the old interface was not significant ($X^2 = 0.0126$, $df = 2$, $p = 0.994$).

Time to success was measured as the time from the participants started the Tor launcher interface to the first successful bootstrap to the Tor network. This measurement was 1) non-normal and heavily right-tailed since participants were less and less likely to succeed as time went on and 2) right-censored at 40 minutes, the maximum time of the experiment. Because of the right-tailed nature of the data, we used a one-tailed Mann–Whitney test. Because the Mann–Whitney test does not account for right-censored data, we assigned the participants who did not succeed the maximum time of 40 minutes for the purpose of this test. We do not find the difference of times to success between participants with the new interface and participants with the old interface to be significant ($Z = -1.84$, $p = 0.0328$, $r = 0.172$).

Active configuration time was measured as the time the participants spent in the Tor launcher interface, except for the time spent the progress screen. For our participants who did not succeed to connect to Tor, subtracted the time spent on the progress screen from the maximum experiment time of 40 minutes. This measurement was also 1) non-normal and heavily right tailed and 2) impacted by right-censored measurement of time to success. Because of the right-tailed nature of the data, we used a one-tailed Mann–Whitney test. We do find the

difference of active configuration times between participants with the new interface and participants with the old interface to be significant ($Z = -3.28$, $p = 0.000516$, $r = 0.307$).

- P118, data not used (Fig. 13)
- P119, data not used (Fig. 14)

J Custom Bridge Attempts

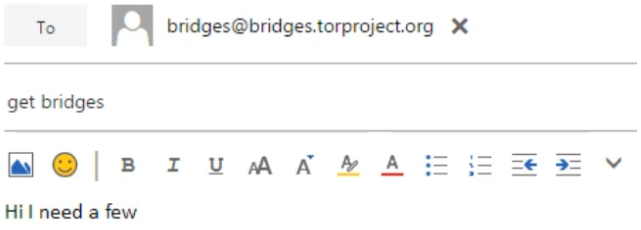


Fig. 11. The message that P95 attempted to send to the bridge auto-responder. The message was sent to the wrong address: bridges@bridges.torproject.org instead of bridges@torproject.org.

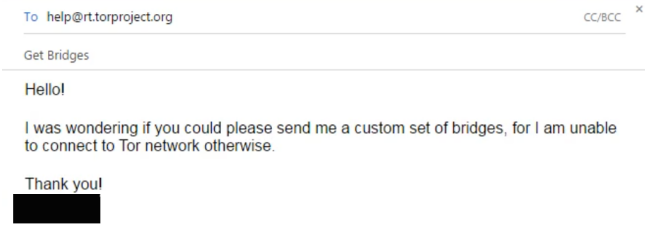


Fig. 12. The message that P110 accidentally sent to the helpdesk rather than the bridge auto-responder. The participant should have used the address bridges@torproject.org, not help@rt.torproject.org. Since the helpdesk address is not an auto-responder, there was no reply.

K Tor Downloads

Participants that downloaded Tor Browser (5/124):

- P115 (Fig. 17)
- P116 (Fig. 16)
- P117, data not used (Fig. 15)

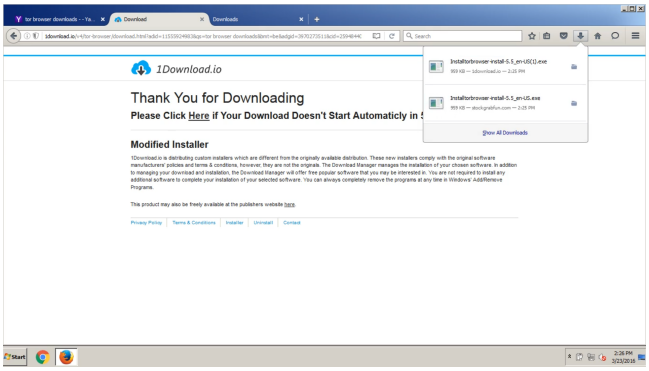


Fig. 13. P118 attempted to download Tor Browser from TechSpot.com, but clicked on a download link which was not from TechSpot.com. The download was from another site, 1download.io.

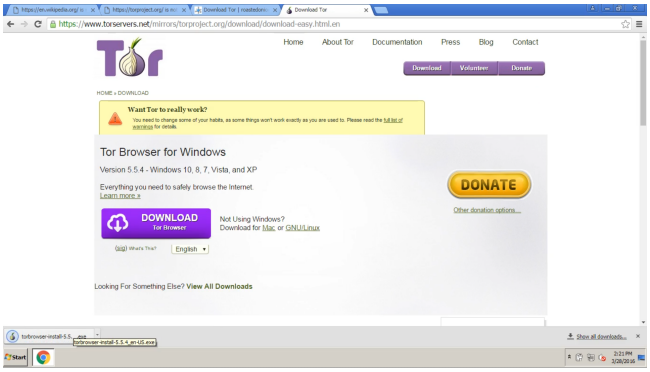


Fig. 15. P117 downloaded Tor Browser from a legitimate Tor Browser mirror, www.torservers.net/mirrors.

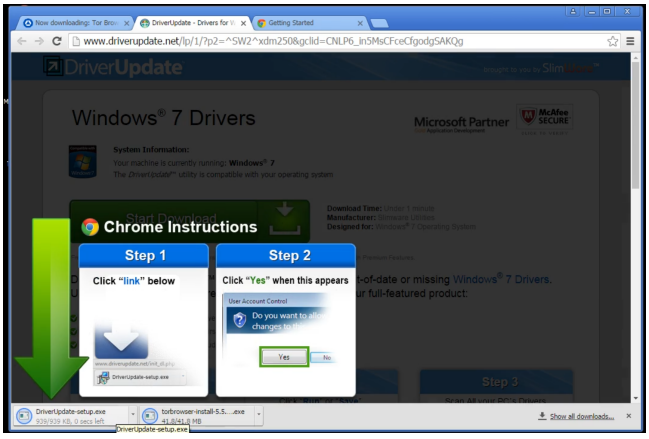


Fig. 14. P119 downloaded Tor Browser from DriverUpdate, a suspicious source. Two executables are downloaded, and the website instructs the user to run an executable named "DriverUpdate-setup.exe," which the participant dutifully does before running the Tor Browser executable.

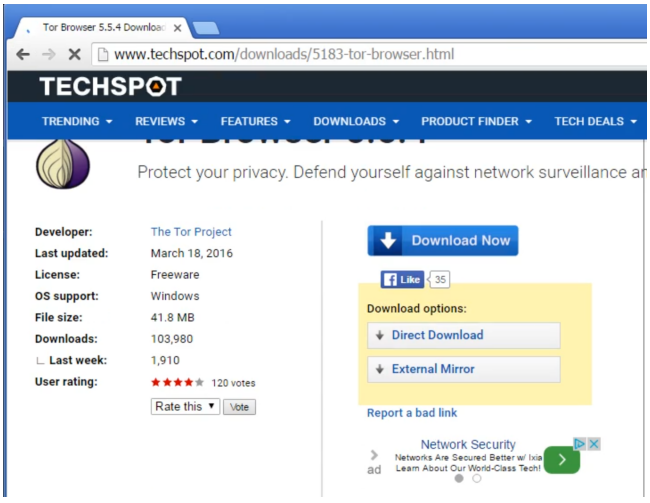


Fig. 16. P116 downloaded Tor Browser from TechSpot, clicking on the download link from TechSpot (the white "Direct Download" button). X10-20160323-132505 in Fig. 13 clicked the blue "Download Now" button and downloaded from a different source.

L All Participant Times

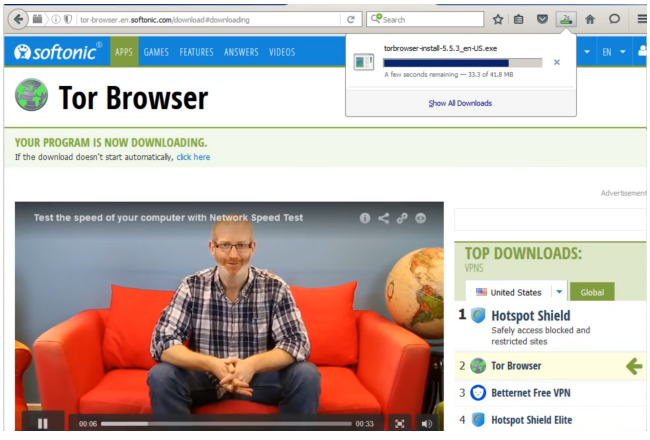


Fig. 17. P115 downloaded Tor Browser from Softsonic, a reputable-looking but unconfirmed site.

		First		Proxy		Bridge		Progress	
E1-NEW	P1	0:03	20%	0:00	0%	0:00	0%	0:10	72%
	P2	0:02	13%	0:00	0%	0:00	0%	0:11	79%
	P3	0:07	51%	0:00	0%	0:00	0%	0:06	42%
	P4	0:06	39%	0:00	0%	0:00	0%	0:08	55%
	P5	0:05	30%	0:00	0%	0:00	0%	0:10	64%
	P6	0:05	28%	0:00	0%	0:00	0%	0:11	63%
	P7	0:04	23%	0:00	0%	0:00	0%	0:13	72%
	P8	0:05	25%	0:00	0%	0:00	0%	0:12	60%
	P9	0:05	23%	0:00	0%	0:00	0%	0:14	71%
	P10	0:08	40%	0:00	0%	0:00	0%	0:11	55%
	P11	0:02	8%	0:00	0%	0:00	0%	0:18	87%
	P12	0:08	39%	0:00	0%	0:00	0%	0:11	54%
	P13	0:08	36%	0:00	0%	0:00	0%	0:12	59%
	P14	0:19	56%	0:00	0%	0:00	0%	0:14	41%
	P15	0:06	16%	0:05	15%	0:06	15%	0:16	45%
	P16	0:03	8%	0:00	0%	0:00	0%	0:33	83%
	P17	0:18	28%	0:00	0%	0:00	0%	0:47	71%
	P18	0:06	8%	0:15	19%	0:30	39%	0:18	23%
	P19	1:14	86%	0:00	0%	0:00	0%	0:11	13%
E1-OLD	P20	0:06	35%	0:00	0%	0:00	0%	0:10	54%
	P21	0:06	30%	0:00	0%	0:00	0%	0:12	64%
	P22	0:12	55%	0:00	0%	0:00	0%	0:09	41%
	P23	0:16	64%	0:00	0%	0:00	0%	0:08	32%
	P24	0:13	50%	0:00	0%	0:00	0%	0:12	46%
	P25	0:16	55%	0:00	0%	0:00	0%	0:09	31%
	P26	0:17	54%	0:00	0%	0:00	0%	0:13	42%
	P27	0:28	64%	0:00	0%	0:00	0%	0:13	29%
	P28	0:09	17%	0:10	19%	0:08	15%	0:22	39%
	P29	0:11	18%	0:26	43%	0:13	21%	0:09	15%
	P30	0:24	36%	0:00	0%	0:00	0%	0:09	14%
	P31	0:15	16%	0:32	35%	0:19	21%	0:24	26%
	P32	0:10	11%	0:45	48%	0:28	30%	0:10	11%
	P33	0:13	11%	0:00	0%	0:00	0%	0:12	10%
	P34	0:11	8%	1:40	70%	0:12	8%	0:20	14%
	P35	0:08	4%	2:17	75%	0:14	8%	0:19	10%
	P36	0:18	9%	0:00	0%	0:00	0%	1:58	63%
	P37	2:25	78%	0:12	7%	0:10	5%	0:18	10%
	P38	0:18	5%	4:01	73%	0:51	15%	0:19	6%
E2-NEW	P39	0:07	19%	0:04	10%	0:09	22%	0:17	42%
	P40	0:06	14%	0:04	11%	0:09	22%	0:19	46%
	P41	0:30	25%	0:46	38%	0:14	11%	0:22	18%
	P42	0:06	4%	0:11	7%	0:12	8%	1:52	75%
	P43	0:11	7%	0:07	5%	0:18	12%	1:31	60%
	P44	1:02	41%	0:08	5%	0:13	8%	0:39	26%
	P45	0:10	6%	0:42	26%	0:48	30%	0:48	30%
	P46	0:05	3%	0:08	5%	0:05	3%	2:17	83%
	P47	0:16	8%	0:09	5%	0:10	5%	2:26	76%
	P48	0:12	6%	0:14	7%	0:08	4%	2:49	80%
	P49	0:07	3%	0:12	5%	0:15	7%	2:57	81%
	P50	0:10	4%	0:06	2%	0:11	5%	3:22	86%
	P51	0:31	11%	0:09	3%	0:17	6%	1:14	26%
	P52	0:10	3%	0:07	2%	0:08	2%	4:58	89%
	P53	0:20	5%	0:14	4%	0:05	1%	5:35	87%
	P54	0:20	3%	0:09	1%	0:09	1%	11:44	94%
	P55	0:25	3%	0:25	3%	0:27	3%	12:22	88%
	P56	0:18	1%	0:11	1%	0:20	1%	23:58	95%
E2-OLD	P58	0:18	32%	0:04	7%	0:08	13%	0:26	45%
	P59	0:10	13%	0:14	18%	0:18	23%	0:34	43%
	P60	0:10	12%	0:04	5%	0:13	15%	0:53	64%
	P61	0:11	10%	1:10	62%	0:13	11%	0:19	16%
	P62	0:17	14%	0:05	4%	0:10	9%	1:26	72%
	P63	0:12	7%	0:47	30%	0:16	10%	1:22	52%
	P64	0:11	7%	0:05	3%	0:14	8%	2:21	82%
	P65	0:22	11%	0:07	4%	0:13	7%	2:27	76%
	P66	0:15	5%	2:26	51%	0:17	6%	1:48	38%
	P67	0:17	6%	1:54	38%	0:36	12%	2:06	42%
	P68	0:39	12%	0:19	6%	2:01	39%	2:00	38%
	P69	0:14	4%	2:11	37%	0:22	6%	1:48	31%
	P70	0:31	7%	1:04	14%	0:28	6%	3:07	41%
	P71	0:10	1%	7:39	69%	0:55	8%	0:20	3%
	P72	0:19	2%	10:26	69%	0:20	2%	3:24	23%
	P73	3:18	10%	10:44	34%	1:01	3%	15:26	49%
	P74	0:16	1%	0:00	0%	0:00	0%	38:54	97%
	P75	0:27	1%	8:60	22%	0:35	1%	29:15	73%
	P76	1:32	4%	0:16	1%	0:20	1%	37:30	93%
E3-NEW	P77	0:18	7%	0:12	5%	0:43	16%	3:01	68%
	P78	1:08	16%	0:20	5%	0:23	6%	4:04	59%
	P79	0:22	5%	0:06	1%	0:26	6%	6:00	82%
	P80	0:16	3%	0:35	7%	0:30	6%	7:14	81%
	P81	0:22	4%	0:26	5%	0:34	6%	7:04	79%
	P82	0:20	3%	1:06	9%	0:53	7%	9:23	77%
	P83	6:00	48%	0:14	2%	0:31	4%	5:38	45%
	P84	0:36	4%	0:15	2%	0:47	5%	13:17	87%
	P85	0:20	2%	0:14	1%	0:27	3%	14:54	84%
	P86	0:48	4%	5:13	26%	0:50	4%	12:52	63%
	P87	0:32	3%	0:14	1%	1:11	6%	11:51	55%
	P88	0:23	2%	1:52	8%	1:22	6%	10:30	45%
	P89	0:49	3%	0:25	1%	1:22	4%	27:57	89%
	P90	0:13	1%	0:03	0%	0:02	0%	39:28	98%
	P92	2:34	6%	3:23	8%	0:53	2%	31:31	79%
	P93	0:19	1%	0:00	0%	0:00	0%	39:10	98%
	P94	5:37	14%	8:34	21%	0:45	2%	22:50	57%
	P95	0:20	1%	0:08	0%	0:07	0%	1:56	5%
	P96	1:19	3%	6:08	15%	6:05	15%	24:13	60%
E3-OLD	P97	0:12	4%	0:16	5%	0:16	5%	4:12	81%
	P98	0:29	8%	0:21	6%	0:21	6%	3:40	58%
	P99	0:41	9%	0:31	7%	1:21	18%	4:38	62%
	P100	0:12	3%	0:41	9%	0:40	9%	6:03	79%
	P101	0:34	5%	0:26	3%	0:59	8%	10:32	83%
	P102	0:40	3%	10:50	45%	1:25	6%	10:50	45%
	P103	0:27	2%	7:06	29%	5:26	22%	11:06	46%
	P104	0:34	2%	1:19	5%	0:49	3%	21:16	87%
	P105	0:44	2%	6:21	20%	6:39	20%	14:50	46%
	P106	0:57	2%	20:54	52%	1:42	4%	15:17	38%
	P107	1:56	5%	7:29	19%	3:37	9%	26:21	66%
	P108	0:21	1%	0:03	0%	0:09	0%	7:36	19%
	P109	1:46	4%	34:25	86%	1:27	4%	2:18	6%
	P110	0:41	2%	5:07	13%	2:23	6%	31:35	79%
	P111	8:46	22%	15:31	39%	8:15	21%	6:25	16%
	P112	0:21	1%	0:15	1%	0:12	0%	38:09	95%
	P113	0:31	1%	0:36	1%	0:15	1%	37:56	95%
	P114	0:50	2%	4:11	10%	8:10	20%	26:38	66%
	P115	0:14	1%	20:10	50%	0:47	2%	18:19	46%
	P116	0:53	2%	7:20	18%	1:26	4%	29:55	75%

Table 5. Times spent on each screen, per participant.