Block Cookies, Not Websites: Analysing Mental Models and Usability of the Privacy-Preserving Browser Extension CookieBlock

Author(s):
Schöni, Lorin; Kubicek, Karel; Zimmermann, Verena

Publication date:
2023

Permanent link:
https://doi.org/10.3929/ethz-b-000638291

Originally published in:
Proceedings on Privacy Enhancing Technologies 2024(1), https://doi.org/10.56553/popets-2024-0012

Block Cookies, Not Websites: Analysing Mental Models and Usability of the Privacy-Preserving Browser Extension CookieBlock

Lorin Schöni
ETH Zürich
Zürich, Switzerland
lorin.schoeni@gess.ethz.ch

Karel Kubicek
ETH Zürich
Zürich, Switzerland
karel.kubicek@inf.ethz.ch

Verena Zimmermann
ETH Zürich
Zürich, Switzerland
verena.zimmermann@gess.ethz.ch

ABSTRACT

In the modern web, users are confronted with a plethora of complex privacy-related decisions about cookies and consent, often compounded by misleading policies and deceptive patterns. Past efforts to enhance online privacy have failed due to their dependence on website compliance. A solution to this lies in privacy-enhancing tools that are directly controlled by the user. However, challenges related to the usability and flawed understanding of the tools' functionality hinder their widespread adoption. To address this problem, we evaluated the browser extension CookieBlock as an example of a current tool, which supports users by blocking tracking cookies independent of website compliance.

We used a complementary approach consisting of an expert evaluation of CookieBlock and the related tools NoScript and Ghostery, and a laboratory user study focusing on the unique details of how users interact with CookieBlock specifically. The laboratory study with 42 participants investigated usage, mental models, and usability of CookieBlock based on eye tracking, interaction, and self-report data. While CookieBlock received good usability ratings, 18 participants were unable to solve a website breakage caused by cookie misclassification on their own. Overall, the results revealed flawed mental models of CookieBlock's functionality and resulting challenges in making the connection between website breakage and cookie misclassification. Implications for CookieBlock and related applications include interface design recommendations supporting accurate mental models and the proposal of improved heuristics to better guide users and warn them about potential identified website breakage.

KEYWORDS

usable privacy, cookies, browser extension, eye tracking

1 INTRODUCTION

Tracking is ubiquitous in the modern web, with the vast majority of websites tracking user data [68, 74]. Recent research indicated that over 90% of websites use cookies to store and track user information [69, 74]. Although cookies have many beneficial effects and are frequently used for website functionality, they are also employed for tracking purposes and personalised advertisements. Major providers, such as Google Analytics, are capable of tracking a substantial portion of a person's overall browsing activities [68]. Many websites implement pervasive tracking through tricks [68], such as using long-lasting cookies with rolling life periods that never expire, even if users were to opt out of tracking in cookie notices [7, 69].

Since the introduction of the EU's GDPR [28] and California's CCPA [46], websites have to provide users with a higher degree of control over how cookies are used. For example, over 62% of European most used websites use cookie notices [21]. However, these and related measures aimed at compliance through legislation did not lead to significant changes in the prevalence of third-party web tracking overall [83]. Furthermore, despite their apparent good intention, consent notices have received criticism for their inadequacy and manipulative strategies. Many of the cookie notices employ deceptive patterns, designed to nudge users towards privacy-invasive settings that may not be in their best interest [30, 51, 62, 73], such as by highlighting an "Accept All" button as compared to the option to reject all. Consequently, rejecting cookies takes substantially more time than simply accepting them [36]. This is further complicated by Consent Management Platforms (CMPs) providing deceptive design options, which are increasingly used by website providers [77, 82]. Further, cookie notices often fail to contain accurate information about the purpose of cookies, creating a mismatch between legal requirements, advertised policy, and actual cookie behaviour. Recent research has consistently shown that the vast majority of websites violate the GDPR [10, 36, 45, 58, 62] or the CCPA [85]. This reality underlines the problems caused by poorly enforced regulations and websites that do not adhere to them.

Most internet users also do not have high technical affinity and lack an in-depth understanding of technical systems or tracking [9, 47, 76]. Therefore, users prefer choices enabling effortless privacy decisions [45]. Furthermore, privacy-preserving behaviour is a secondary goal of users, as they primarily want to complete tasks of greater personal importance, such as getting work done, connecting with other people, or online shopping. Even if people had the technical understanding and intention to browse the internet privately, this might not transfer into actual behaviour. This effect has been described as a privacy paradox, where users do not take any substantial steps to protect their online privacy despite holding strong privacy beliefs [5, 9].

Privacy-enhancing technologies (PETs) such as browser extensions have great potential to assist especially lay users in preserving their privacy without consuming attentional resources or detracting from primary goals. As such, users can complete tasks relevant to their primary goal while a tool takes care of secondary privacy

Proceedings on Privacy Enhancing Technologies 2024(1), 192–216

https://doi.org/10.56553/popets-2024-0012

concerns. However, if the tool is overly complex, users may not be able to effectively utilise it to enhance their privacy. Moreover, such tools might require a high amount of input or interaction and can end up consuming more attentional resources than they save, thereby defeating their purpose. For example, the NoScript browser extension [53] offers a variety of powerful tools to block most forms of tracking, but often ends up disrupting functionality of websites [2] before site-specific exceptions have been made, and is therefore not a suitable solution for the average user. Furthermore, if a tool does not work or is not used as intended based on a user's lack of understanding, it could even compromise user privacy. Accordingly, the design of PETs requires a careful balance between ease of use and its provided features.

With regard to PETs addressing the challenge of tracking ubiquity, developers have proposed privacy-enhancing browser extensions that prevent privacy intrusion on the client side, independently of a website's functionality and without performance losses [11]. These extension are either ad blockers (e.g., uBlock Origin [35], AdBlock [1, 65]) blocking both advertising and tracking, or they specialise on tracking only (e.g., Ghostery [24], PrivacyBadger [27], DuckDuckGo Privacy Essentials [23]). These browser extensions directly interfere with a user's web browser and are not reliant on website compliance. This is especially relevant as the factors influencing website providers' decisions rarely include the users' privacy [83]. Therefore, tools working independent of website compliance are some of the most popular choices for privacy-enhancing tools and widely adopted [54, 58, 76]. However, they only block web requests and do not interfere with the cookies used. A current browser extension addressing that challenge is CookieBlock that we further describe next.

1.1 Background: CookieBlock

CookieBlock is a browser extension proposed by Bollinger et al. [10] that fills the gap of other privacy extensions by removing tracking cookies in a flexible way. Rather than relying on lists, it uses machine learning to classify cookies according to their purpose as strictly-necessary, functional, analytics, or advertising, categories defined by ICC UK Guide [39]. Users can select which categories they reject and cookies classified as such are then removed while browsing. Fig. 1 illustrates a snapshot from the features of the installed extension in the browser. As reported by the authors, the model achieves 84.4% balanced accuracy, comparable to human expertise. Nevertheless, when the model misclassifies necessary cookies, their removal can break website functionality. Bollinger et al. [10, Sec. 5.3] evaluated breakage on 100 websites, finding that CookieBlock caused eight websites to forget users' setting (causing consent notice to reappear or disabling the language selection), and seven websites had their registration or login functionality broken. The authors thus implemented an option to add exceptions for given domains or pause cookie removal. These interruptions are part of a larger problem affecting PETs to counteract tracking on websites, as due to the largely modular nature of these sites, privacy-intrusive functions are often embedded in larger systems that are generally useful for users [72].

For our research, we selected CookieBlock as a prime example of current privacy-enhancing browser extensions, which are still facing technical challenges with regard to website breakage that may affect usability, users' understanding, and thus actual use. Our research has two goals: First, we aim to better understand the usability challenges faced by CookieBlock in comparison to related tools with either different technical functionality (e.g., static algorithm vs. machine learning approach) or a slightly different focus (e.g., blocking all tracking vs. only cookie-related tracking technologies). Second, we aim to explore users' mental models of CookieBlock, how users interact with the browser extension, and how users evaluate CookieBlock. These insights can then be extended to related PETs that can affect usability in general, and will provide more information on how users approach usability issues in PETs and how they could be overcome.

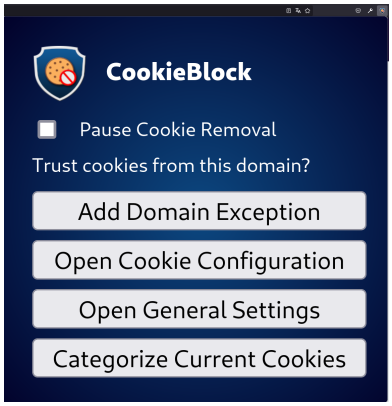


Figure 1: CookieBlock's features after installation.

1.2 Research Questions & Contribution

To understand these questions from different viewpoints, we used a complementary approach consisting of an expert evaluation and a laboratory study with users. The heuristic experts evaluation adds an expert perspective and focuses on a high-level comparison between existing tools. We compared the machine-learning based tool CookieBlock and the related tools NoScript and Ghostery.

As outlined in the introduction, NoScript is a powerful but also disruptive tool based on a static algorithm, thereby potentially leading to a high privacy-usability trade-off. In comparison, Ghostery is based on curated blocklists and heuristic elements, making it less disruptive and potentially resulting in a lower privacy-usability trade-off with unlikely website breakage. Both tools are actively maintained, with Ghostery having a particularly large user base (see Table 4 in Appendix C for a comparison). In the expert evaluation we investigated the following research question:

- **RQ1**: **How does the privacy-usability trade-off differ between privacy-related browser extensions that differ with regard to their aims, functionality and technical implementation?** We would like to analyse potential usability challenges resulting from these differences and potential website breakage.

The laboratory study adds a user perspective and provides an in-depth analysis of the challenges encountered when users interact with CookieBlock, especially focusing on potential website

breakage and supporting users in coping with it. In the laboratory study, we investigated the following research questions:

- **RQ2**: **What are the users' mental models of CookieBlock?** We would like to analyse the users' understanding of CookieBlock's privacy-preserving functionality, its settings and especially its role in website breakage due to misclassifications.
- **RQ3**: **How do users interact with CookieBlock?** We aim to explore which elements users focus on, whether users understand when a website breakage is caused by CookieBlock, and how they solve the problem.
- **RQ4**: **How do users evaluate CookieBlock?** We are interested in how users rate their experience with and the usability of the extension, what problems they perceive, and what suggestions for improvement they provide.

To answer these questions, we conducted a heuristic evaluation [60] with N=4 experts as well as a user study in which N=42 participants installed, configured, and used CookieBlock in two tasks. We combined eye tracking data with other behavioural measures, demographic information, ratings from standardised scales like the System Usability Scale, and qualitative analysis of open-ended questions. We first break down the results individually before synthesising them to gain a comprehensive understanding of how users experience CookieBlock. Through this multi-faceted approach, we identify weaknesses in the current design and show that participants' mental model is often inaccurate, ultimately leaving 18 participants unable to solve a website breakage caused by cookie misclassification on their own. Finally, we discuss consequences of our analysis, implement proposals to clarify expected functionality and improve the interface design, and propose a heuristic to target the main usability challenge of CookieBlock resulting from potential cookie misclassifications.

Our research provides the following contributions:

- We describe and reflect on a surprising gap between the implementation of CookieBlock as a promising browser extension aiming to support users in protecting their privacy across websites and the users' often inaccurate mental models. We further explore challenges users encounter when the extension leads to the breakage of a website. Therefore, we study not only the usability of the browser extension working as intended, but explicitly look into the user interaction and mental models of users facing a website breakage caused by CookieBlock.
- Extending related work, we describe the findings of a user study combining previously used self-report and interaction data with additional behavioural eye tracking data that provides valuable insights into attention, mental workload, and mental models. The data shows that mental workload is especially high when coping with a broken website and that the challenge does not lie in the users' interaction with the extension but in making the connection between the broken website and the extension.
- From the findings, we derive implications and design recommendations for the further development of CookieBlock, similar tools, and related research in Section 6.1. We make the suggestions available in a pull request[1] for easy implementation and evaluation with CookieBlock or related tools in future work.

2 RELATED WORK

After exploring previous work on consent notices, we examine related work on privacy-related browser extensions and their use.

2.1 Consent Notices & Dark Patterns

Cookie consent notices are intended to determine which tracking users agree to. However, most of these consent notices use deceptive patterns [31, 69, 73, 84]. These patterns are intended to nudge users towards less privacy-friendly choices, such as by making 'reject all' options less visible or removing them altogether. Accordingly, they violate GDPR requirements, which for example demand equal visibility for accept and reject options. Deceptive pattern designs consistently increase tracking acceptance [6, 51, 52, 84] by abusing user's quick, automated thinking [12]. The prevalence of deceptive patterns is further increased through templates supplied by content management providers (CMPs), of which over 90% use deceptive patterns [62, 84]. However, there are also large differences between regions, with primarily only EU and US websites taking steps to inform users of privacy choices, likely due to GDPR compliance [69]. In the context of smartphone applications, Kollnig et al. [44] even found that consent is often not obtained before engaging in third party tracking thereby potentially violating current privacy regulations.

In addition, even choices expressed on a cookie consent notice might simply have no effect. Sanchez-Rola et al. [69] found that 90% of sites employ long-lasting tracking cookies no matter what the user has chosen. Consent notices can also negatively affect the user experience more directly. Habib et al. [34] found that the impact of consent notices on usability varies greatly based on design decisions, but always proves to be a barrier to users thereby reducing website usability. These results indicate that consent notices are an inadequate solution even with the backing of well-known legislation, as they rely on the compliance of websites.

2.2 Browser Extensions

Various tools were introduced to counteract tracking without apparent success. The World Wide Web Consortium (*W3C*) proposed the Platform for Privacy Preferences (*P3P*), a standard for websites to transmit their privacy policies, which could then be read by browsers [19]. However, as Leon et al. [48] demonstrated, many websites simply bypassed these recommended policies by misrepresenting data to browsers. Later, W3C introduced a Do Not Track (*DNT*) header [71], which informed sites of user preferences so they can adjust their tracking. Roesner et al. [68] analysed anti-tracking defence mechanisms and found that blocking third-party cookies and using DNT is insufficient for privacy protection. Following these difficulties and lack of industry adoption, the DNT project was eventually discontinued [16]. DNT is now going through a reincarnation as Global Privacy Control [33], this time backed legally by California's CCPA. However, these past experiences show the weakness of mechanisms that rely on website compliance. To overcome

[1] https://github.com/dibollinger/CookieBlock/pull/15

this challenge, browser extensions are promising tools that can operate independently of specific websites to enhance a user's privacy by either showing additional information or automatically dealing with consent notices based on predefined settings. Prominent examples include Consent-O-Matic [61], Cliqz Autoconsent [49], CookieEnforcer [43] and CookieGlasses [55]. However, these browser extensions still have shortcomings, as they ultimately rely on website compliance to implement the cookie notice as intended without any additional tracking. In contrast, we evaluate the CookieBlock browser extension for our study, which automatically blocks tracking cookies and does not rely on website compliance. Table 4 in Appendix C illustrates the main differences in the aims, functionalities, and the underlying technology or method of currently available privacy-related browser extensions to provide a quick overview and comparison with CookieBlock.

Besides technical aspects, previous work on privacy-related browser extensions has focused on the users' mental models, their interaction with, and evaluation of those tools.

Mental Models. In a survey with 48 participants, Cranor et al. [20] studied online behaviour regarding advertising and privacy tools after users watched informational videos about advertising and tracking. Reception to advertising was mixed, with some welcoming targeted ads, while other participants were surprised and scared by the degree of tracking that websites employ. Participants were also interested in tools to counteract tracking but found most privacy-enhancing tools to either be too basic in their instructions, or too technical with jargon. Therefore, the authors highlighted the need for a carefully balanced middle ground. Another key takeaway was that participants wanted protection tools, but crucially did not want those tools to interfere with website functionality. Our study builds on Cranor et al.'s study [20] by evaluating CookieBlock as a tool that offers the functionality to counteract tracking that Cranor et al.'s participants expressed interest in. We explore whether it meets the suggested "middle ground" with regard to understanding and expected non-interference with website functionality.

Furthermore, Schaub et al. [70] examined changes in privacy concerns and awareness of people using different anti-tracking browser extensions. Their study indicated that users' privacy awareness is inconsistent and inaccurate, oftentimes overestimating or underestimating the extent with which companies track users and collect information about them. The authors also concluded that the design of browser extensions is crucial in helping users to get better mental models of tracking and privacy in the internet. Mathur et al. [54] explored how people use a variety of browser extensions that blocked ads and compared users' mental models to the actual functionality of the extensions. They found that most users' mental models are underdeveloped even if they use these extensions, suggesting that awareness of the deeper functions is generally limited. These findings indicate that users' mental models of privacy-related browser extensions are largely influenced by the extension's design and can in turn significantly influence the interaction with it [47, 54]. Therefore, we explore the users' mental models of CookieBlock to analyse this relationship further and to extend related work on user mental models.

Interaction & Evaluation. Hubert et al. [38] compared the usability of Adblock, uBlock, Ghostery, and Privacy Badger by giving participants tasks to use them on a number of real-world websites and found large inconsistencies and differences in usability. Notably, no browser extension managed to meet high usability thresholds. Leon et al. [47] investigated how users approach privacy-preserving browser extensions like Adblock Plus or Ghostery. In a laboratory study, 45 participants were assigned to install, configure, and finally use one specific browser extension for five tasks, three of which could malfunction based on which browser extension was active. Afterwards, the tools were evaluated with the System Usability Scale (SUS) [14]. Despite coming from an educated background, none of the participants could initially demonstrate an understanding of web tracking mechanisms. Furthermore, while installation seemed easy for most participants, tool configuration was generally seen as difficult. Overall, all tools scored between 40 and 50 out of 100 points on the SUS, indicating low usability [3]. The authors concluded with recommendations to specifically improve communication and feedback that does not rely on technical jargon. Finally, they also noted that participants had difficulties understanding that problems with websites were caused by the tools they are using.

Corner et al. [17] conducted a laboratory study with 30 participants in the UK and tasked them with using DuckDuckGo Privacy Essentials, Ghostery, and Privacy Badger browser extensions, together with think-aloud and completing the SUS. Overall, the usability scores were substantially higher compared to Leon et al. [47], with ratings of 60, 79, and 62, respectively. However, the difference is likely caused by the task design that focused on exploring the extension's functionality, instead of the extension interfering with some other goal. Participants still made various suggestions for improvement, such as improved visual feedback, easier access to help resources, as well as real-time feedback on the extension's activities. Similar to [17, 47], we include self-report data and the SUS to compare usability metrics with previous work. However, in contrast we complement the collection of self-report and behavioural interaction measures with eye tracking data. As eye tracking measures a continuous stream of data that is indicative of a user's visual attention, it becomes possible to determine how much time and effort is spent on any specific element. For example, we can determine how much the interaction with CookieBlock detracts from primary goals such as the interaction with the website.

Contribution. To summarise, our study addresses the open research challenge to support users in protecting their privacy when handling tracking cookies. To overcome the problems associated with website compliance, we analyse CookieBlock as a promising browser extension that does not rely on website compliance. We build on previous work by including comparable mental models, interaction, and usability measures to be able to derive implications for related tools based on comparisons. However, our work differs from other work in that we specifically study a challenging website breakage scenario and additionally include eye tracking as a promising technology for usability studies. We report on the findings of a complementary expert evaluation and user study, and the lessons learned that can inform future work.

3 EXPERT EVALUATION

In an expert evaluation, N=4 experts with a background in psychology and human-computer interaction analysed the usability of CookieBlock in comparison to two similar browser extensions via a

heuristic evaluation [60]. The heuristic evaluation was chosen as it has been demonstrated to find the most usability problems as compared to other techniques while having a good cost-benefit ratio [42]. On purpose, the experts were chosen with a background in human sciences and usability as compared to technical disciplines to evaluate the lay user perspective rather than technical aspects. The number of experts followed recommendations by Nielsen [59].

We compared the browser extension CookieBlock with two tools selected from a larger pool of previously researched tracker blocker or ad blocker browser extensions in the privacy context. The tools NoScript and Ghostery were chosen from the category of tools with tracker blocking as their primary purpose, in contrast to other tools that block ads or handle consent notices, to increase comparability to CookieBlock (see Appendix C). As all tracker blocker tools were based on different technologies and differed in terms of functionality, we were interested in tools that (a) primarily differed in their anticipated privacy-usability trade-off and (b) tools that were at the time available and applicable in the same way as CookieBlock, i.e., being installable from the Chrome web store. Therefore, we selected the tools NoScript and Ghostery that are based on a static algorithm and curated blocklists, respectively. They are both actively maintained and used by a large number of users. As illustrated in Section 2.2, this prominence has led both NoScript [2] and Ghostery [17, 38, 47] to be investigated in previous usability studies which enables comparisons with related work. However, they vary in their level of disruptiveness, which potentially affects the privacy-usability trade-off. The selected tools were thus not only relevant to study from a user perspective but also did not rely on other software, such as DuckDuckGo Privacy Essentials that only works in conjunction with the DuckDuckGo search engine. NoScript offers many options for blocking JavaScript, which is used by most forms of tracking. Based on a static algorithm, it offers a suite of tools for further customization. However, its blocking often ends up disrupting the functionality of websites [2]. Ghostery focuses on blocking trackers based on a curated blocklist but also automatically blocks cookies through consent notices, making it much less likely to impair site functionality.

3.1 Heuristic Evaluation Procedure

The process of the evaluation was as follows: *First,* even though the experts had familiarized with the ten principles used for the heuristic evaluation proposed by Nielsen ([40, 60], cf. Table 5), the principles were again explained to ensure a common understanding. For example, these principles included visibility of system status, consistency, and error prevention. *Second,* the three tools CookieBlock, NoScript and Ghostery were analysed. Each tool was first installed by the experts in the Chrome web browser (Installation), then the available settings were inspected (Settings). Finally, the interaction with and effect of the tool was analysed in use with exemplary real websites (Use). For each aspect - Installation, Settings and Use - the ten principles were applied. To make sure that no aspect was overlooked, each expert considered all ten principles. Yet, as considering all principles at the same time can be challenging, each expert was in the lead for three select principles. During the evaluation, a joint protocol of all findings was created and checked by all experts to ensure completeness and to avoid misunderstandings. Screenshot examples of the different GUIs are shown in Fig. 7 to facilitate understanding of the findings. *Third,* the experts exchanged final conclusions after having analysed all three tools. All findings collected in the joint protocol for each tool, aspect, and principle are shown in Tables 6 to 8 within Appendix C. In the following, we discuss the main findings and implications.

3.2 Heuristic Evaluation Findings

CookieBlock offers immediate feedback upon installation, but suffers from a lack of separation between options, as well as unclear feedback on their purpose. The settings allow for in-depth customization, but rely on advanced technical terms and lack visual cues. During use, while the extension provides multiple options to address breakage, they are not clearly communicated and impede error recovery. Furthermore, the extension does not clearly communicate its actions to the user.

NoScript lacks feedback after installation, further complicated by discrepancies between the store logo and the extension icon. In the settings menu, it could benefit from clearer explanations of certain terms, such as "privileged site" and options concerning default, trusted, and untrusted settings. However, it does utilize tabs for easy overview and separates advanced settings from general settings, allowing for a more organized experience. During use, symbols in the interface are not always intuitive (see Fig. 7), and there is a lack of feedback on specific actions such as what happens if the lock symbol is clicked. When content is blocked, the NoScript symbol on the blocked content allows for making a connection between blocking or breakage and Noscript. However, there is no clear instructions on what to do or how to recover.

Ghostery impresses with a visually appealing installation process and consistent symbols, as seen in Fig. 7. The settings menu utilizes clear standards such as color-coding and underlined text for links and information symbols leading to more information. An indication of currently blocked trackers and a statistical overview over past activity are particularly informative. However, it is unclear whether users can make the link between potential website breakage and Ghostery's actions. Finally, some of the advanced options like submitting your own tracker are challenging to understand.

The findings highlight that the three tools all offer distinct benefits but ultimately differ substantially in their purpose and target user groups. CookieBlock and NoScript both offer advanced and customizable blocking features, but their design is less intuitive and relies on expert knowledge. However, while CookieBlock limits blocking to cookies only, NoScript has a much broader impact that affects more aspects of a web page. Ghostery on the other hand follows a more minimalist and aesthetic design and appears relatively undisruptive with regard to website functionality, but also offers limited depth and flexibility in its features.

4 USER STUDY METHOD

In a laboratory study, 42 participants installed and used the CookieBlock browser extension in two web shop scenarios. In those scenarios, they encountered two situations in a randomised order, one where CookieBlock worked as intended and one where CookieBlock broke the website. In the breaking condition, the log-in

functionality failed. The study adopted a within-subject design, with participants interacting with CookieBlock in the same way, but with a randomised sequence of working and breaking conditions to balance effects. We analysed the participants' interaction with CookieBlock, their expectations before and after the study, as well as the perceived usability of the tool across scenarios. Participants used Google Chrome for the study, as it was the most widely used web browser worldwide [75].

In the following sections, we first detail the study procedure, the implemented web shops used to test different CookieBlock use cases, the eye tracking technology used to explore user interaction and attention. Afterwards, we describe the sample, discuss ethical considerations, and highlight our analysis.

4.1 Procedure

To provide an overview, the procedure is summarised in Fig. 2. At the start of the study, participants were presented with an informed consent sheet and any questions were clarified. If they chose to proceed, they first filled out a background questionnaire, which included Affinity for Technology Interaction (ATI) [29] and Internet Users' Information Privacy Concerns (IUIPC) [32] scales. Afterwards, we calibrated the eye tracker with a standard 9-point calibration procedure. We then asked participants to read CookieBlock's description in the Google Chrome web store and to install the extension on the lab PC so that they experienced a realistic set-up process. Participants then filled out a questionnaire about their experiences and understanding of CookieBlock to explore mental models, as well as the System Usability Scale (SUS) [14].

In the main part of the study, participants were twice instructed to buy tickets for a specific event on our web shops. To ensure that all participants had an identical environment that was not affected by their behaviour during the installation task, for this step we loaded a different Chrome profile with CookieBlock pre-installed, pinned to the taskbar, and with standard settings. In the working condition, CookieBlock did not interfere and they could proceed as usual. In the breaking condition, CookieBlock intentionally misclassified a login cookie and blocked it, which caused the login to fail and prevented participants from purchasing tickets. Accordingly, participants had to anticipate that CookieBlock caused the disruption and use CookieBlock's options to resolve the problem or disable the extension altogether. If participants were not able to make any progress within five minutes, they were told that "the CookieBlock browser extension causes the problem." If they were still not able to login within two more minutes, the task was skipped and we proceeded to the next step. After each condition, participants filled out additional questionnaires regarding usability and positive and negative experiences of using CookieBlock.

The final part of the study consisted of a questionnaire that again asked about the users' understanding of CookieBlock to evaluate a potential impact of using the extension on the users' mental model. It also asked for participants' feedback on potential design improvements. The entire procedure took around 30 to 40 minutes.

4.2 Task Environment

We created two fictional web ticket shops to emulate basic variants that participants could encounter during browsing. To increase realism, the web shops were interactive and contained descriptions and images for several events. In the breaking condition, they also provided feedback similar to what people would encounter on a real website. These web shops were hosted locally and not connected to the Internet. Both web shops shared basic functionality but only differed slightly in their design so that they could be differentiated. Participants were instructed to first identify a specific event (like a concert) on the web shop and then to buy tickets for that event. In order to purchase the tickets, participants needed to login, for which they received login credentials on a task sheet. Screenshots of both web shops are given in Appendix B.

4.3 Eye Tracking

To understand how people use PETs, eye tracking serves as a good measure that can enhance questionnaires and other sources of behavioural data [64]. As eye tracking measures a continuous stream of data that is indicative of a user's visual attention [37], it becomes possible to determine how much time and effort is spent on any specific task or visual area, and therefore how much it detracts from any primary goal.

We used a Tobii Pro Fusion eye tracker [80] with standard properties: a sample rate of 250 Hz, an accuracy of 0.3° and a precision of 0.04°. The tracker was mounted at the bottom of the computer monitor. The monitor used was a Dell S2522HG with a resolution of 1920x1080 and a size of 24.5 inches. Participants were seated approximately 65 cm from the monitor. We used Tobii Pro Lab v1.207 [81] to record and process all eye tracking data from both eyes. Before collecting eye tracking data, we calibrated the tracker using a standard 9-point calibration. Following best practices [15, 37] and since we were not interested in very small elements like text characters, the calibration was accepted if average accuracy and precision errors were both < 1.00° of visual angle. All eye tracking measures were collected within roughly 20 minutes after calibration ended, so no drift corrections were taken.

In eye tracking, fixation refers to times where the eyes are focused on a specific point and relatively stationary, typically for smaller periods of time below one second. Saccades are rapid eye movements that occur between periods of fixation, allowing the eyes to move quickly from one point to another, such as when moving between words during reading. We segmented fixations and saccades in Tobii Pro Lab with standard settings using a velocity threshold of 30°/s, meaning that periods above that speed were classified as saccades.

4.4 Sample

A total of 42 participants took part in our study. Two were excluded because they were not able to install CookieBlock on their own, leaving us with a final sample of N=40 participants. Of these, 17 identified as female, 22 as male, and one as diverse. The participants' age distribution was as follows: 19 were between 18-24, 14 between 25-34, seven between 35-44. 22 participants have a university degree and 18 participants completed secondary education. 16 participants indicated they had used a browser extension related to privacy before, while 24 stated they did not. The sample's privacy concerns measured with the IUIPC-8 ranging from "1 - strongly disagree" to "7 - strongly agree" was M=5.84 (SD=0.87) for the scale Control,

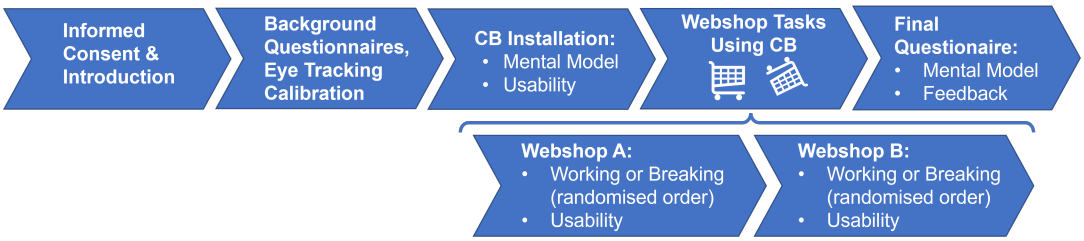


Figure 2: Visual summary of the study procedure.

M=6.46 (SD=0.75) for the scale Awareness and M=5.45 (SD= 1.04) for the scale Collection. The means and variance are similar to a larger pool representative of the UK population [32] (see Table 9 for a comparison), which suggests our sample has normal privacy concerns. The sample's affinity for technology interaction measured with the ATI scale ranging from "1 - completely disagree" to "6 - completely agree" was M=3.81 (SD=0.93), slightly higher than the average score of 3.5 found in the general population [29].

4.5 Ethical Considerations

The study design followed established ethical guidelines for psychological research involving humans [18] and was approved by our university's ethics board. The approved ethics application stated that individual data is stored securely, may only be accessed within the university for research purposes, and cannot be released publicly. We minimised the potential for privacy-invasion by collecting only vague demographic information, such as using age range brackets. The eye tracking data did not contain any images of faces or eyes; we merely collected data about the corresponding coordinate points on the computer screen that the participant's eyes focused on.

Participants were recruited from a voluntary opt-in database associated with the university that consists not only of students but people of different age groups and occupations. Through the database, participants were already informed about the nature of the tasks and that the study involved eye tracking. Before the study, participants were provided an informed consent sheet describing the details and rights related to data collection and storage. Participation was voluntary and participants could abort the study at any time without negative consequences. Furthermore, participants could request the deletion of their data at any time, without giving any reason or suffering negative consequences. All participants received an equal payment that was calculated based on the hourly wages for student assistants. Finally, participants received an additional background information sheet with contact details after the conclusion of their session.

4.6 Analysis

For the quantitative analysis, we used t-tests or one-way ANOVAs to analyse significant differences between two or more groups, respectively, such as whether task speed differs significantly between people that require help and people that do not. However, if the test assumptions were not met, e.g., due to non-parametric data, we instead used more robust procedures such as the Wilcoxon rank sum test [13]. Finally, we used Fisher's exact test to evaluate categorical differences between multiple levels. This test was used in place of a χ^2 test due to the low number of observations in multiple cells of the contingency table (cf. Table 2 in [78]). Using the Benjamini-Hochberg procedure [8] we corrected p-values for tests concerning the same variables within each research question to limit the possibility of false positive results due to multiple testing [79].

For the qualitative analysis, two independent raters used a deductive approach [56] to code mental models and negative consequences accuracy following pre-set categories with prototypical examples in a codebook. For classifying accuracy, we relied on the functionality description provided by CookieBlock, i.e. "CookieBlock allows you to automatically remove cookies that collect sensitive data and track you. It uses machine learning to automatically predict the purpose that each cookie is used for." The initial inter-rater agreement was 76%. For other open-ended responses, such as design suggestions, two raters instead used open coding to cluster comments into categories.

5 USER STUDY RESULTS

The following sections summarise the findings structured by our three research questions that guided the laboratory user study.

5.1 RQ2: Mental Models

We first analysed users' mental models of CookieBlock and its functionality by asking participants to describe their understanding of CookieBlock's functionality twice, once after the installation but before actual use (*Before*) and once after having interacted with CookieBlock (*After*). Furthermore, prior to starting the scenarios, we asked participants to provide a prediction of potential negative consequences related to using CookieBlock. The textual answers were categorized in terms of accuracy as "accurate," "partially accurate" and "not accurate" following a deductive coding approach [56]. The results are shown in Table 1 and illustrated in the following: An example for accurate answers was *"CookieBlock uses machine learning to identify and block cookies on websites" (P13).* Partially accurate answers contained phrases like *"Presumably, the extension sees which cookies a website uses and prevents them from functioning" (P26).* And finally, answers classified as not accurate contained statements such as *"The extension looks at cookie pop-up windows, recognizes these, and automatically changes cookie preferences to the intended settings" (P12).*

For each participant, the answer before and after interacting with CookieBlock was compared to explore influences on the mental

model. We found that mental model accuracy improved in nine cases, decreased in nine cases, and did not change in 19 cases[2]. Seven participants with increased accuracy gained the understanding that CookieBlock blocks cookies in the background and does not interact with consent notices. On the other hand, participants with decreased accuracy mentioned confusion related to website breakage, speculating whether CookieBlock might simply block all cookies or whether websites can recognize the extension and therefore cause an intentional breakage.

Table 1: Mental Model Accuracy

Item	Inaccurate	Partially accurate	Accurate
Before	15	11	14
After	14	12	12
Prediction	20	10	8

Most participants had a flawed understanding of potential negative consequences using CookieBlock could bring with it (*Prediction*, see Table 1). Only eight participants could foresee that it could interfere with website functionality (even though stated in the installation description). 10 participants correctly anticipated a disruption of sorts but could not correctly foresee how that would manifest itself, such as guessing that it might *"break storage of passwords" (P42)*. A total of 20 participants could not foresee any negative consequences or made completely inaccurate guesses. Of these, seven participants were concerned about data collection from CookieBlock itself, thinking that instead of the websites, CookieBlock could now track and sell their browsing data.

We tested whether mental model accuracy or the prediction accuracy of negative consequences could predict task success in the following scenarios. Based on the sample size and its unequal cell distribution, we used Fisher's exact test to confirm whether there was a significant effect. We found that mental model accuracy was not significantly related to task success ($p = .127$), but prediction accuracy was ($p = .046$). Furthermore, an ANOVA revealed that mental model accuracy ($F(2, 37) = 1.14, p = .331$) and prediction accuracy ($F(2, 35) = 3.32, p = .096$) do not have a significant effect on time required to solve the task.

5.2 RQ3: User Interaction with CookieBlock

Overall, 22 participants were able to solve the breaking condition on their own within five minutes. However, 18 participants were unable to login within that time and required a prompt after five minutes, informing them that CookieBlock was the cause of the breakage. We limited any intervention to this single prompt so that we could still observe how people approached the task without steering their action towards a particular CookieBlock option or skipping the task. Naturally, participants that did not require assistance (median of 3.49 minutes) were much faster ($t(33.65) = 5.19, p < .001, d = -1.66$) to complete the task compared to those that required assistance (median of 5.57 minutes).

A total of 25 participants used the intended standard procedure of resolving the breakage by opening the CookieBlock drop-down menu and either pausing cookie removal or whitelisting the domain. However, 12 participants (of which nine did not receive any assistance) attempted to solve the breakage in other ways, primarily by opening the settings. One person solved the problem by deactivating the CookieBlock extension itself.

In the Google Chrome browser, extensions are not automatically pinned to the taskbar once they are installed. However, it is important for CookieBlock to remain visible as its options can become highly relevant in case of cookie misclassification as tested in the breaking condition. Accordingly, we investigated how many people pinned the extension after it had been added to Google Chrome. Ultimately, only three people pinned the extension during that time so that the CookieBlock icon, which opens the drop-down menu, was visible. While a relevant finding, it did not affect our other results as users purposefully started the task conditions with CookieBlock pinned.

5.2.1 Eye Tracking. Overall, five participants were excluded from eye tracking data analysis; four did not meet the minimum calibration threshold (average accuracy and precision < 1.00° of visual angle) and one person moved outside the measured range for a prolonged period of time. As such, we analysed eye tracking data from 35 participants.

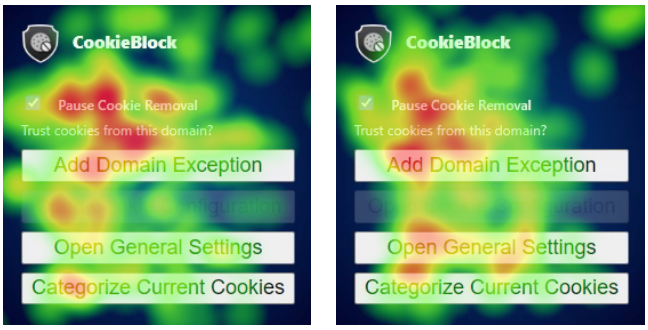
Our eye tracking results clearly indicated that the breaking condition created an environment in which cognitive load— indicative of how much demand is placed on mental resources in a task—was high when compared to the working condition or while participants installed the extension. This was indicated by decreases in saccadic peak velocity [25], increases in pupil dilation [66], and an increased number of saccades per second [87]. Table 10 in Appendix E summarises the purpose of the different measures based on [26, 37, 64]. The detailed eye tracking results are further outlined in Table 11 in Table 10.

In the breaking condition, the median time from first fixation of the CookieBlock icon pinned to the browser to the first mouse click on it was 4.85 seconds. The median time from the first fixation of the drop-down menu that opens once the icon has been clicked to clicking on an option within the menu was only 2.49 seconds. This illustrates that participants hesitated much longer when looking at the icon and came to a much quicker decision on the drop-down menu. However, the clarity of the two elements seems reversed, as the drop-down menu demanded more visual attention with a total fixation duration of 4.36 seconds compared to the icon with only 0.48 seconds. Overall, these results indicate that the icon was easy to process but not quickly understood as relevant to the breakage, while the drop-down menu had a clearer immediate purpose that took more attentional resources to fully grasp.

For participants that required assistance, the total median time to the first fixation was 1.96 minutes for the icon and 3.00 minutes for the drop-down. However, participants that did not require assistance only required median times of 1.15 minutes for the icon and 1.16 minutes for the drop-down. These differences indicate that participants who required assistance spotted the CookieBlock icon long before using it. However, they only fixated on it for very short times and waited long before clicking on it in order to open the drop-down menu. In contrast, participants that did not need assistance seemed to have instantly interacted with the CookieBlock

[2]Due to NAs in some answers, the total number is lower than 40.

icon once spotted and then proceeded to investigate the drop-down menu. These findings could be explained by the fact that participants who required assistance simply were not aware that the CookieBlock icon could provide them with a solution and thus only spent little visual attention on it.



Required help **No help required**

Figure 3: Heat maps of eye fixations on the drop-down menu for people who did vs, did not require help in making the connection between website breakage and CookieBlock.

Apart from analysing the eye tracking data quantitatively, we also conducted a qualitative analysis to compare behaviour immediately after participants clicked the CookieBlock icon and looked at the drop-down menu. Participants that did not require help scanned the drop-down menu more globally and considered more options. In comparison, participants that did require help focused more on direct solutions. This is consistent with our behavioural analysis, which showed that most people who opened the settings did not receive any prior help.

We also investigated whether age or technical affinity could explain whether people were able to solve the breaking task on their own. While the descriptive data revealed a trend of older participants being more likely to solve the problem on their own, a Fisher's exact test could not detect a significant relation between requiring help and the age brackets ($p = .076$). Further, a t-test revealed that technical affinity, measured through the ATI scale, was not significantly related to requiring help either ($t(36.10) = 0.18, p = .853$).

5.3 RQ4: User Evaluation of CookieBlock

Following the installation and two task conditions, we administered the System Usability Scale (SUS) [14] to compare usability evaluations of CookieBlock across the installation task, the working scenario, and the breaking scenario. After installation, the mean SUS scores were $M = 75.81$ ($SD = 15.52$) and after the working condition the mean score was $M = 74.62$ ($SD = 17.92$) out of 100, both indicating good usability [3]. However, after the breaking condition, SUS scores dropped to $M = 61.35$ ($SD = 20.27$), indicating only average usability [3] and below the average score of web applications of 68 [4]. The installation and working scores were similar to SUS scores found for Ghostery with an average of 79. In contrast, SUS scores assigned to DuckDuckGo Privacy Essentials and Privacy Badger in previous work were lower, with respective scores of 60 and 62 [17]. Overall, the scores were consistently higher compared to SUS scores between 40 and 50 in previous research that compared a wide range of tools in a situation where they partially interfered withtasks [47]. Table 2 provides an overview of these values overall and for participants who solved the breaking condition without help, compared to those who needed assistance. Notably, people who required help did not assign lower SUS scores compared to those that did ($z = 233, p = .220$). The average SUS score in the breaking condition of participants who were not able to solve the problem on their own was still relatively high with $M = 58.19$ ($SD = 18.82$).

Table 2: Comparison of SUS Values across Tasks

Group	Installation	Normal Functioning	Breaking
Overall	75.81	74.62	61.35
Accurate models	79.29	75.36	68.46
No support	77.27	76.19	64.05

Opinions. Furthermore, participants were queried for their opinion at each stage of the experiment in an open text field. After installation, nine participants mentioned they were happy with the information provided. However, eight participants criticised the volume of information, either deeming it too much or too little. Half of them advocated for a simpler design with less text, such as stating that *"the instructions are slightly too much to read" (P19).* The other half advocated for a more detailed design, such as by including additional information on *"why the extension is superior or different to other extensions" (P16).* A common theme after installation was also a wish for more visual information, primarily in the form of graphical elements rather than text, which is consistent with recent research indicating that recognizable figures are preferred in comparison to text [86].

After the breaking condition, four participants mentioned using the extension was laborious, while four other participants were satisfied with the extension's usability. 19 participants overall mentioned that they did not fully understand why CookieBlock interfered with the login attempt. Four participants mentioned that breakages should be explained more clearly, so that users would know what to do when one occurs.

After the working condition, opinions were more positive, with 12 participants mentioning good usability and four participants stating that CookieBlock is easy to use once users are familiar with it. These answers are exemplified by P34 stating that *"CookieBlock works exactly like Adblock, which makes using it intuitive and easy"* and P02 mentioning *"it works once you understand what exactly you have installed."* However, four more participants stated that they were uncertain if CookieBlock even did anything.

Overall, participants consistently indicated their confusion at the appearance of cookie consent notices, incorrectly assuming that CookieBlock would automatically deal with it. Five participants specifically noted that the appearance of consent notices in the working condition was a negative experience.

Recommendations. In addition, participants provided retrospective recommendations for CookieBlock after completing all tasks in an open text field. Main themes concerned the provision of visual information or improved information. Nine participants mentioned they desired information when CookieBlock blocks something, such as in the form of a popup to notify the user. Seven participants suggested added statistics that inform them about the number and categories of current and historical cookie blockages. Similar to the opinions stated after the install condition, six participants proposed an improved explanation that better prepares them for the possibility of breakage and how to address such cases. For example, P13 mentioned the options in the drop-down menu might not be intuitive for most users and should be better described. Finally, several participants suggested individual design and noticeability improvements, for example using more notable colours or more clearly separated buttons.

A smaller group of participants made suggestions to enhance CookieBlock's functionality. Five participants suggested an improved algorithm that would substantially reduce cases of misclassification. Two people advocated for a temporary pause feature rather than a permanent pause. P09 suggested that CookieBlock should remove cookie notices when blocking cookies.

Improved Design. After gathering the participants' feedback, we also presented them with a design mock-up we prepared in advance, which added more information to CookieBlock's drop-down menu and icon (see Fig. 4 in Appendix A). Specifically, we proposed a number on the CookieBlock icon, showing users how many cookies are blocked on the current page. In addition, the drop-down menu contained additional statistics about current and past cookie blockages for each category. 15 participants commented that they were happy with the design mock-up while two participants commented that the changes were unnecessary. Four participants suggested that the statistics be shortened or moved to another place so that the drop-down does not appear too large. Finally, seven participants recommended additional design improvements by either increasing colour contrast, making the blue colour brighter, or changing the buttons to be more round and colourful.

Cookie Categories. Participants rated the clarity of the four cookie category descriptions used for classification and based on UK ICC cookie categories [39]: Strictly-Necessary, Functionality, Analytics, and Advertising/Tracking. Ranging from "0 - not clear at all" to "10 - very clear", the overall clarity was rated as M=8.32 (SD=1.74). Additionally, participants were prompted to indicate any unclear categories. The Functionality category was by far mentioned the most, with seven participants mentioning that they did not understand it and five more participants indicating that they did not understand any of the categories. Analytics, Strictly-Necessary, and Advertising/Tracking were mentioned to be unclear by four, three, and one participant, respectively.

6 DISCUSSION & OUTLOOK

In this section, we discuss the findings of the heuristic evaluation representing an expert perspective and the laboratory study focusing on the user perspective concerning the four research questions, compare them with each other and with related work, and derive implications for CookieBlock and related research. However, not all findings of the heuristic evaluation can be directly transferred to CookieBlock due to technology differences. We implemented the majority of the proposed design changes for future work and submitted them as an anonymous pull request accessible along with further explanations at https://anonymous.4open.science/pr/09F2. We then reflect on insights and lessons learned that are relevant for related research beyond the specific extension.

6.1 Summary of Findings and Derivation of Design Recommendations

RQ1: Security-Usability Trade-Off: The findings of the heuristic evaluation show that three tools with expected different security-usability trade-offs — CookieBlock, NoScript and Ghostery — differ substantially in their provided functionalities and settings, even though all share a primary focus on tracker blocking. For example, CookieBlock solely focuses on blocking cookies, provides advanced settings for this feature, but also requires expert knowledge to make use of them. In contrast, NoScript blocks a much broader range of content but thereby also affects website functionality to a larger extent. NoScript users also require expert knowledge, e.g., to enable or disable specific content. In comparison, Ghostery's target group does not necessarily require expert knowledge as it very rarely impacts website functionality, options are explained within the tool, and understanding is supported through design features such as consistent color-coding.

RQ2: Mental Models: Consistent with prior findings [54], mental models of participants were flawed. A third of all users had an incorrect understanding of CookieBlock's functionality (Table 1). Similarly, 19 participants stated they did not fully understand why CookieBlock interfered with the login attempt. We also found that the participants' understanding of CookieBlock's functionality seldom changed after using the extension, but even decreased for 9 participants and only improved for another 9 participants. While somewhat surprising, the finding may be explained by the extension's main purpose, that is, to automatically categorise and block cookies. Therefore, the classification process and its consequences are hidden from the user unless a website breaks. Furthermore, the heuristic evaluation revealed that CookieBlock's classification options were unclear and lacked feedback. Accordingly, it is difficult for many users to understand the classification process and to make the connection between site breakage and CookieBlock even though they were informed about that possibility during installation.

The possibility of misclassifications cannot be excluded completely and the aim regarding user mental models is not to make all users technical experts with detailed knowledge of the classification process. Nevertheless, the development of correct mental models should be supported by interface design, such as the clearer and simpler text choices used to describe the Ghostery browser extension. This may enable users to use CookieBlock according to their privacy preferences and to more easily overcome challenges such as website breakage. The browser store description determines the expectation of CookieBlock's functionality. Since the users with incorrect mental models were also those who failed to realise that CookieBlock was breaking website functionality, better explaining

the functionality could significantly improve the usability. To do so, we propose the following recommendation (R):

- *R1: Improve Extension Description. We suggest an adapted extension description in a short bullet list with emphasis on the key functionalities that were unclear to users.*

Participants also demonstrated an incorrect understanding that CookieBlock itself would track their activity. This might have partially been caused by an option during setup, which asked users for their consent before the extension stores a local history of cookies. This illustrates how important it is for descriptions to be as clear as possible to avoid minimize the risk of misunderstandings.

RQ3: User Interaction with CookieBlock. About half of the participants required help in making the connection between website breakage and CookieBlock. Comparing that aspect with other tools, the expert study showed that even though NoScript impacts website functionality to a larger degree than CookieBlock, in some cases it might at least be easier to make the connection between breakage and NoScript as actively blocked content often carries the NoScript logo on it. Yet, this approach cannot be easily transferred to CookieBlock due to the technology, i.e., actively blocking JavaScript vs. blocking cookies through machine learning.

Of course, it would be ideal for CookieBlock if no website breakage occurred. Yet, misclassification can never be fully avoided in machine learning. Thus, it remains difficult for users to notice that a website's functionality is disrupted by CookieBlock—a persistent challenge to usability. Accordingly, It is still important to minimize the impact or frequency of breakages, as most users prioritise functionality over privacy [20, 67]. Even though CookieBlock cannot perfectly predict functionality disruption, we propose an added heuristic that would notify users about potential website breakage similar to a recently proposed system by Smith et al. [72]. A secondary mechanism predicts when filter lists might break websites. However, since blocked cookies generally lead to issues only after interaction, it is not sufficient to build a classifier based on web crawling. Accordingly, more complex machine learning mechanisms are needed to directly classify occurrences dynamically based on heuristics. This is particularly pertinent to CookieBlock and other PETs that enhance detection through machine learning, but even browser extensions relying on curated lists like Ghostery can in rare cases impact website functionality or content and suffer from similar problems.

CookieBlock could also track repeated submissions of the same form and then ask the user whether functionality of the website seems disrupted. If the user agrees, CookieBlock could disable the removal of involved cookies. Otherwise, the warning would be suppressed for the domain. Yet, such a heuristic is prone to both false positives and negatives. False positives stem from forms supposed to be submitted multiple times, such as search forms, or when a user enters an incorrect password into a login form. In both cases, a simple removal of an advertising cookie would trigger the user notification. False negatives are possible when a necessary cookie, whose creation is independent of a form, is misclassified and removed. For instance, language switching is often implemented through a cookie being set using an HTTP(S) request for a localised address link. To avoid overly high false positive rates and annoying notifications, we recommend tuning such a heuristic towards fewer notifications and therefore low false negative rate (i.e., high precision).

- *R2: Add Secondary Mechanisms.* We recommend secondary mechanisms to identify situations where CookieBlock might interfere with user goals, i.e. through website breakage.

The eye tracking measures indicate that participants were able to process the pinned CookieBlock icon itself and needed to look at it only for a short time, but that they waited much longer before actually clicking it. This difference could be explained by participants not understanding that the icon is relevant for the interruption in the breaking condition. This assumption is reinforced when comparing participants that needed help with those that did not. Those that did not need help quickly looked at the drop-down after fixating on the icon initially, indicating an immediate transition from one element to the other. Participants that did require help had a much longer delay of roughly 17 seconds on average. This shows that understanding the icon's use was crucial for participants to solve the problem on their own, as those who did not do so within five minutes were much less likely to make the connection. In the heuristic evaluation, Ghostery's choice of directly describing the icon in the installation process was noted as a positive heuristic, whereas inconsistent icons in the NoScript extension created confusion for the usability experts. Meaningful design of icons and hints towards their purpose seem to be important factors in enhancing the tool's visibility.

- *R3: Improve Icon Visibility.* The icon should be visible, understandable, and described in the installation process. The icon's visibility should change depending on whether CookieBlock is active, is blocking anything, or is inactive.

Several participants were having trouble understanding the functionality of the "pausing the cookie removal" checkbox and the "add domain exception" button. As noted in the heuristic evaluation, the purpose of these options is not clearly communicated and can only be inferred from their title. While the expert analysis found the use of symbols to be consistent, they alone may not be sufficient to describe options as users' knowledge of internal processes is low. Accordingly, it is important to add additional information, especially considering that many users might have inaccurate mental models that do not match up with CookieBlock's actual functionality.

- *R4: Add Explanatory Tooltips.* We recommend tooltips explaining the interface and visually distinguishing the interface for resolving website breakage from other settings.

RQ4: User Evaluation of CookieBlock. Participants had the option to evaluate the extension and to provide feedback on design suggestions. The feedback indicated that while cookie categories were clear, the fact that CookieBlock itself does not remove cookie notices was unclear to many participants. This problem could either be solved by providing more clarification to users, e.g., with regard to suitable extensions accomplishing this task in addition to CookieBlock, or by adapting CookieBlock to block consent notices as well. Another solution could come in the form of information about the extension's activity in the form of statistical overviews. The use of statistics for the Ghostery extension was noted as a positive feature allowing deeper insights into past activity. Furthermore, the

display of statistics would directly inform users about differences between the categories based on the amount of blocked cookies.

- *R5: Add Statistics.* We recommend adding statistics on removed cookies based on our participants' suggestions. This might provide feedback on whether the extension is active and on what happens in the background.
- *R6a: Clarify Non-Blocking of Consent Notices.* We suggest explicitly stating that cookie notices are not removed by CookieBlock and adding links to notice-removing extensions. We also suggest modifying the statement about local cookie history to counteract the incorrect assumption that CookieBlock itself tracks users.
- *R6b: Block Consent Notices.* Alternatively, CookieBlock could expand its features to also block cookie consent notices.

Participants provided mixed feedback as to the extent of information they would like to see, with some advocating for more while others suggested less text. Accordingly, we recommend a more modular concise approach, with an option for users to get more insights in a separate element. This approach mirrors the design choice of Ghostery, which tries to minimize upfront text and options so that users can focus on a few but important elements with fewer distractions. While the scope of options and descriptions for CookieBlock and NoScript were extensive, such a level of detail might not be understandable for the average user and would take too much attentional load to process.

- *R7: Design Concise Descriptions.* We recommend descriptions to be concise without relying on overly technical language. Further information can be put on a separate element people could access through a button.

Cookie categories received a high mean clarity rating of 8.32. However, the functionality category seemed to be unclear with seven people indicating uncertainty, and five more expressing confusion at all categories. Notably, the category system of many other PETs is different, many of which use their own classification. For example, Ghostery uses the category "Social Media", which provides information about its source, but not about its purpose. On the other hand, some PETs like NoScript simply do not assign any categories. Considering this, the ICC UK cookie categories [39] adapted for CookieBlock seem to be feasible descriptions that could be used by other PETs, yet could still be improved with regard to the functionality category.

- *R8: Use Clear Categories* We recommend PETs to employ clear classification categories that follow established standards, like the ICC UK categories for cookies [39]. These should be evaluated to further increase understandability.

Many users wished for a popup, indicating that cookies are blocked. Previous research has also indicated that users rely on direct prompts to gain awareness [67]. However, such a feature is unlikely to be well-received for a tool that is often active, as many pages would trigger a popup. This would likely overstimulate users and quickly become more of a nuisance than a help. It is possible that participants simply underestimated the frequency by which cookies are set and therefore blocked, or that they only focused on how to solve the problem they just encountered without considering additional downsides. The design changes we proposed to participants in the questionnaire (Fig. 4 in Appendix A) might be a less annoying alternative, as there is a simple static number indicating the number of currently blocked cookies.

6.2 Eye Tracking

Eye tracking for evaluating user interaction with a privacy-enhancing technology proved highly beneficial with regard to triangulation. The eye tracking data provided us with valuable behavioural measures that could be used to explain some self-report data but also to provide additional insights. For example, eye tracking provided us with an indication of high mental workload when dealing with website breakage that was not visible from the SUS data alone. Additionally, we gained insights into participant's visual attention. From the fixation data, we observed that participants had more difficulty connecting website breakage to CookieBlock than interacting with the extension itself. From that, we could derive important design recommendations. Furthermore, the measure was non-intrusive and did not negatively impact the participants' options to interact with the computer and study tasks.

Relevant lessons learned for related studies include the challenge to set the area of interest for very small objects such as small icons in the browser. Furthermore, there are challenges when comparing dynamic content on websites, such as tracking gaze and fixation over pages that include scrolling. For example, it might be a good idea to isolate specific UI elements and test their noticeability in a more controlled setting before moving on to environments that more closely resemble real-world conditions. Furthermore, even though we carefully controlled experimental conditions like lighting, distance to the screen, or wearing glasses, we still had to exclude five participants despite repeated calibration attempts. For statistical analyses requiring a certain sample size, it would thus be recommended to plan with additional participants beyond considering drop-out for other reasons.

6.3 Implications for Related & Future Research

While the direct implications for CookieBlock have already been discussed in Section 6.1, the following Table 3 provides a summary of identified usability issues and recommendations aimed at addressing them. From this, we derive implications on a) how these recommendations could be applied to other PETs being affected by similar issues, as well as b) suggestions for further research into their effects in similar contexts or with similar study designs.

Future work could investigate advances in AI technology and their effect on algorithms of tools like CookieBlock, which have a promising future to become more commonplace and more potent. However, a major barrier to adoption of PETs with strong privacy protection is their potential to negatively affect usability. To address this, the impact of such downsides should be resolved or decreased. One major improvement to CookieBlock would be a method that detects breakages, such as blocking a login. For example, it could recognise important forms like login prompts, and then provide a visual cue if submission does not work due to a misclassified cookie. While this method would be complex to implement, it could enhance usability by substantial degrees.

Two participants failed to install CookieBlock. While the number is too limited for any conclusion, both were at least 55 years old.

Table 3: Summary of Recommendations for CookieBlock and Derived Implications for Related Research

Usability Issues	CookieBlock Recommendation	Implications for Related PETs	Implications for Related Research
User understanding is flawed	R1: Improve Extension Description	Make users aware of potential negative consequences, how to identify, and handle them	Explore how mental models translate from an initial setup text to active use
Misclassification is hard to spot	R2: Add Secondary Mechanisms	Make use of additional mechanisms to assist the user experience in case of PETs' adverse impact	Explore how users could be assisted with challenges stemming from PETs use
Users are unaware of icon	R3: Improve Icon Visibility	Make the first point of interaction as clear as possible, e.g., by highlighting it during setup	Explore unobtrusive yet notable design options in web browsers
Options are unclear to users	R4: Add Explanatory Tooltips	Tooltips can provide additional information when concise descriptions (see R7) are hard to understand	Explore usability and understanding of information provided in tooltips
Activity is not visible to users	R5: Add Statistics	Statistics can provide unintrusive feedback on the tool's functionality to the user	Explore the potential of statistics to support accurate mental models
Users misunderstand complex functionality	R6: Clarify Non-Blocking of Consent Notices or Block Consent Notices	PETs developers should weigh focusing on one functionality while addressing limitations against including more functionalities that increase complexity	Explore user perceptions and expectations with regard to these trade-offs
Users dislike long descriptions	R7: Design Concise Descriptions	Keep descriptions concise and add additional information to other optional elements	Explore how information volume differences at different stages of PETs use affect mental models
Users need clear categories	R8: Keep using clear Categories	Use generally feasible categories and description, e.g., those of UK ICC	Analyse potential to improve understandability of categories (especially functionality) further

Further studies should consider the technical affinity of older people and their ability to not only use but also setup PETs. We have also observed a tendency of younger participants being less likely to solve the problem on their own compared to older participants. This is surprising, considering we would expect opposite effects due to younger people growing up with digital technology and thus being more proficient at using it. While this effect might be due to a limited sample size, future work could investigate age differences in web browser extension proficiency.

6.4 Limitations

Study Design. As outlined by McGrath [57], researchers face the dilemma that they cannot maximise all dimensions relevant for research validity. These comprise a) generalizability to other populations, b) precision in control and measurement, and c) realism of the study context. In our study, we aimed for a highly controlled laboratory setting, which among others are important for the accuracy of eye tracking data. We aimed to enhance realism by having participants follow a realistic installation process and interacting with CookieBlock in a web shop setting. Yet, for example, the breaking condition impacting log-in functionality occurs much less frequently in actual browsing (ca. 8% of cases [10]).

Mental Models. Our questions asking for the understanding of CookieBlock indicates the participants' mental models, e.g., by showing a relation between accuracy and task performance. However, we can not directly infer understanding from the responses as they might be confounded by reading and memorising (as compared to "understanding") the text from the installation process (even for the second questionnaire shown 15-30 minutes later). For an in-depth analysis of mental models, future work should thus deploy a suitable knowledge test or other option to demonstrate and explain understanding. Previous research has used drawings to infer mental models [63], but such an approach is time-consuming and adds additional variability when analysing the results. Accordingly, it is important to strike a good balance and implement a test that is not too taxing on participants' time or mental resources.

Sample. Our user study sample was of limited diversity as, due to the laboratory setting, all participants were recruited in the same country. Furthermore, most participants were students. Two participants from the age group over 55 were not able to install CookieBlock and thus excluded from further analysis. Yet, our age variation was too limited to further explore how very large age differences affect both understanding and using CookieBlock.

7 CONCLUSION

We evaluated how users approach the privacy-preserving browser extension CookieBlock that protects users from tracking by selectively blocking cookies. We compared the extension with two other tools in a heuristic expert evaluation to illustrate usability differences. In addition, we analysed user interaction in a work-as-intended situation and in a situation where the extension directly interfered with the primary task by breaking website functionality. Based on our findings, we make several suggestions with implications for CookieBlock, PETs in general, as well as research efforts in related areas. Since almost half of the participants were unable to solve the breaking condition on their own, the privacy-usability trade-off seems to be a major barrier. Thus, a focus should be on making PETs easy and intuitive to use, and to explain the situations in which users have to interact with PETs.

Data Availability Statement. The data that support the findings of this article are openly available at https://doi.org/10.3929/ethz-b-000627400.

ACKNOWLEDGMENTS

We thank Yanis Isenring and Linda Fanconi for their contribution to data collection. We also thank Linda Fanconi for contributions to analysis. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] AdBlock. 2023. AdBlock - What can we help you with? https://helpcenter.getadblock.com/hc/en-us; Accessed on 2023.08.20.

[2] Abdul Haddi Amjad, Zubair Shafiq, and Muhammad Ali Gulzar. 2023. Blocking JavaScript without Breaking the Web: An Empirical Investigation. http://arxiv.org/abs/2302.01182 arXiv:2302.01182 [cs].

[3] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies* 4, 3 (2009), 114–123.

[4] Aaron Bangor, Philip T. Kortum, and James T. Miller. 2008. An Empirical Evaluation of the System Usability Scale. *International Journal of Human–Computer Interaction* 24, 6 (July 2008), 574–594. https://doi.org/10.1080/10447310802205776

[5] Susanne Barth and Menno D.T. de Jong. 2017. The Privacy Paradox Investigating Discrepancies between Expressed Privacy Concerns and Actual Online Behavior A Systematic Literature Review. *Telemat. Inf.* 34, 7 (Nov. 2017), 1038–1058. https://doi.org/10.1016/j.tele.2017.04.013 Place: USA Publisher: Pergamon Press, Inc..

[6] Jan M. Bauer, Regitze Bergstrøm, and Rune Foss-Madsen. 2021. Are you sure, you want a cookie? – The effects of choice architecture on users' decisions about sharing private online data. *Computers in Human Behavior* 120 (July 2021), 106729. https://doi.org/10.1016/j.chb.2021.106729

[7] Paschalis Bekos, Panagiotis Papadopoulos, Evangelos P. Markatos, and Nicolas Kourtellis. 2022. The Hitchhiker's Guide to Facebook Web Tracking with Invisible Pixels and Click IDs. http://arxiv.org/abs/2208.00710 arXiv:2208.00710 [cs].

[8] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57, 1 (1995), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2517-6161.1995.tb02031.x.

[9] Sophie C. Boerman, Sanne Kruikemeier, and Frederik J. Zuiderveen Borgesius. 2021. Exploring Motivations for Online Privacy Protection Behavior: Insights From Panel Data. *Communication Research* 48, 7 (Oct. 2021), 953–977. http://journals.sagepub.com/doi/10.1177/0093650218800915

[10] Dino Bollinger, Karel Kubicek, Carlos Cotrini, and David Basin. 2022. Automating Cookie Consent and GDPR Violation Detection. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, MA, 2893–2910. https://www.usenix.org/conference/usenixsecurity22/presentation/bollinger

[11] Kevin Borgolte and Nick Feamster. 2020. Understanding the Performance Costs and Benefits of Privacy-focused Browser Extensions. In *Proceedings of The Web Conference 2020*. ACM, Taipei Taiwan, 2275–2286. https://doi.org/10.1145/3366423.3380292

[12] C. Bösch, B. Erb, F. Kargl, Henning Kopp, and Stefan Pfattheicher. 2016. Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns. *Proceedings on Privacy Enhancing Technologies* 2016 (2016), 237 – 254.

[13] Patrick D Bridge and Shlomo S Sawilowsky. 1999. Increasing Physicians' Awareness of the Impact of Statistics on Research Outcomes: Comparative Power of the t-test and Wilcoxon Rank-Sum Test in Small Samples Applied Research. *Journal of Clinical Epidemiology* 52, 3 (1999), 229–235. https://doi.org/10.1016/S0895-4356(98)00168-1

[14] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[15] Benjamin T. Carter and Steven G. Luke. 2020. Best practices in eye tracking research. *International Journal of Psychophysiology* 155 (Sept. 2020), 49–62. https://doi.org/10.1016/j.ijpsycho.2020.05.010

[16] W3C contributors. 2019. "WG closed – w3c/dnt@5d85d6c". https://github.com/w3c/dnt/commit/5d85d6c3; Accessed on 2023.02.28.

[17] Matthew Corner, Huseyin Dogan, Alexios Mylonas, and Francis Djabri. 2019. A Usability Evaluation of Privacy Add-ons for Web Browsers. In *Design, User Experience, and Usability. Practice and Case Studies*. Vol. 11586. Springer International Publishing, Cham, 442–458. https://doi.org/10.1007/978-3-030-23535-2_33 Series Title: Lecture Notes in Computer Science.

[18] Council for International Organizations of Medical Sciences (CIOMS). 2016. *International Ethical Guidelines for Health-related Research involving Humans*. Technical Report. Council for International Organizations of Medical Sciences (CIOMS). https://doi.org/10.56759/rgxl7405

[19] Lorrie Faith Cranor. 2003. P3P: Making Privacy Policies More Useful. *IEEE Security and Privacy* 1, 6 (Nov. 2003), 50–55. https://doi.org/10.1109/MSECP.2003.1253568 https://www.w3.org/TR/P3P11/.

[20] Lorrie Faith Cranor. 2012. Can users control online behavioral advertising effectively? *IEEE Security & Privacy* 10, 2 (2012), 93–96.

[21] Martin Degeling, Christine Utz, Christopher Lentzsch, Henry Hosseini, Florian Schaub, and Thorsten Holz. 2018. We Value Your Privacy... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy. In *Proceedings of the 26th Annual Network and Distributed System Security Symposium (NDSS 2019)*. Network and Distributed Systems Security (NDSS) Symposium 2019, San Diego, CA, USA, 1–15. https://doi.org/10.14722/ndss.2019.23378

[22] Scott DeVaney. 2021. uBlock Origin—everything you need to know about the ad blocker. https://addons.mozilla.org/blog/ublock-origin-everything-you-need-to-know-about-the-ad-blocker/; Accessed 2023.08.20.

[23] Developers of DuckDuckGo. 2023. DuckDuckGo Privacy Essentials. https://github.com/duckduckgo/duckduckgo-privacy-extension; Accessed on 2023.02.13.

[24] Developers of Ghostery GmbH. 2023. Ghostery: Best Ad Blocker & Privacy Browser. https://www.ghostery.com/; Accessed on 2023.02.13.

[25] Leandro L. Di Stasi, Rebekka Renner, Peggy Staehr, Jens R. Helmert, Boris M. Velichkovsky, José J. Cañas, Andrés Catena, and Sebastian Pannasch. 2010. Saccadic Peak Velocity Sensitivity to Variations in Mental Workload. *Aviation, Space, and Environmental Medicine* 81, 4 (April 2010), 413–417. https://doi.org/10.3357/ASEM.2579.2010

[26] Claudia Ehmke and Stephanie Wilson. 2007. Identifying web usability problems from eye-tracking data. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but not as we know it - Volume 1 (BCS-HCI '07)*. BCS Learning & Development Ltd., Swindon, UK, 119–128.

[27] Electronic Frontier Foundation. 2019. Privacy Badger. https://privacybadger.org/; Accessed on 2023.02.28.

[28] European Parliament, Council of the European Union. 2016. Regulation (EU) 2016/679 Of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). http://data.europa.eu/eli/reg/2016/679/2016-05-04; Last accessed on: 2023.02.28.

[29] Thomas Franke, Christiane Attig, and Daniel Wessel. 2019. A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human–Computer Interaction* 35, 6 (2019), 456–467.

[30] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174108

[31] Paul Graßl, Hanna Schraffenberger, Frederik Zuiderveen Borgesius, and Moniek Buijzen. 2021. Dark and Bright Patterns in Cookie Consent Requests. *Journal of Digital Social Research* 3 (Feb. 2021), 1–38. https://doi.org/10.33621/jdsr.v3i1.54

[32] Thomas Groß. 2021. Validity and reliability of the scale internet users' information privacy concerns (IUIPC). *Proceedings on Privacy Enhancing Technologies* 2021, 2 (2021), 235–258. https://doi.org/10.2478/popets-2021-0026

[33] GPC group. 2023. Global Privacy Control (GPC). https://globalprivacycontrol.org/.

[34] Hana Habib, Megan Li, Ellie Young, and Lorrie Cranor. 2022. "Okay, whatever": An Evaluation of Cookie Consent Interfaces. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–27. https://doi.org/10.1145/3491102.3501985

[35] Raymond Hill. 2023. uBlock Origin. https://github.com/gorhill/uBlock; Accessed on 2023.01.05.

[36] Maximilian Hils, Daniel W. Woods, and Rainer Böhme. 2020. Measuring the Emergence of Consent Management on the Web. In *Proceedings of the ACM Internet Measurement Conference*. ACM, Virtual Event USA, 317–332. https://doi.org/10.1145/3419394.3423647

[37] Kenneth Holmqvist, Marcus Nystrom, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost Van De Weijer. 2015. *Eye Tracking*. Oxford University Press, London, England.

[38] Marvin Hubert, Joachim Griesbaum, and Christa Womser-Hacker. 2020. Usability von Browsererweiterungen zum Schutz vor Tracking. *Information - Wissenschaft & Praxis* 71, 2-3 (April 2020), 95–106. https://doi.org/10.1515/iwp-2020-2075 Publisher: De Gruyter Saur Section: Information – Wissenschaft & Praxis.

[39] International Chamber of Commerce UK. 2012. ICC UK Cookie guide. https://www.cookielaw.org/wp-content/uploads/2019/12/icc_uk_cookiesguide_revnov.pdf; Accessed on 2023.02.28.

[40] Jakob Nielsen. 2020. 10 Usability Heuristics for User Interface Design. https://www.nngroup.com/articles/ten-usability-heuristics/.

[41] Rolf Bagge Janus Bager Kristensen. 2020. Consent-O-Matic. https://github.com/cavi-au/Consent-O-Matic.

[42] Robin Jeffries, James R Miller, Cathleen Wharton, and Kathy Uyeda. 1991. User interface evaluation in the real world: a comparison of four techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, New York, NY, USA, 119–124.

[43] Rishabh Khandelwal, Asmit Nayak, Hamza Harkous, and Kassem Fawaz. 2022. CookieEnforcer: Automated Cookie Notice Analysis and Enforcement. https://doi.org/10.48550/ARXIV.2204.04221

[44] Konrad Kollnig, Reuben Binns, Pierre Dewitte, Max Van Kleek, Ge Wang, Daniel Omeiza, Helena Webb, and Nigel Shadbolt. 2023. A Fait Accompli? An Empirical Study into the Absence of Consent to Third-Party Tracking in Android Apps. In *Proceedings of the Seventeenth USENIX Conference on Usable Privacy and Security (SOUPS'21)*. USENIX Association, USA, Article 10, 15 pages.

[45] Michael Kretschmer, Jan Pennekamp, and Klaus Wehrle. 2021. Cookie Banners and Privacy Policies: Measuring the Impact of the GDPR on the Web. *ACM Transactions on the Web* 15, 4 (Nov. 2021), 1–42. https://doi.org/10.1145/3466722

[46] California Legislature. 2018. California Consumer Privacy Act of 2018. https://www.oag.ca.gov/privacy/ccpa

[47] Pedro Leon, Blase Ur, Richard Shay, Yang Wang, Rebecca Balebako, and Lorrie Cranor. 2012. Why Johnny can't opt out: a usability evaluation of tools to limit online behavioral advertising. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin Texas USA, 589–598. https://doi.org/10.1145/2207676.2207759

[48] Pedro Giovanni Leon, Lorrie Faith Cranor, Aleecia M. McDonald, and Robert McGuire. 2010. Token attempt: the misrepresentation of website privacy policies through the misuse of p3p compact policy tokens. In *Proceedings of the 9th annual ACM workshop on Privacy in the electronic society*. ACM, Chicago Illinois USA, 93–104. https://doi.org/10.1145/1866919.1866932

[49] Sam Macbeth. 2020. Cliqz Autoconsent. https://github.com/cliqz-oss/autoconsent.

[50] Sam Macbeth. 2022. Autoconsent. https://addons.mozilla.org/en-US/firefox/addon/autoconsent/; Accessed 2023.08.20.

[51] Dominique Machuletz and Rainer Böhme. 2020. Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR. *Proceedings on Privacy Enhancing Technologies* 2020, 2 (apr 2020), 481–498. https://doi.org/10.2478/popets-2020-0037

[52] Stefan Mager and Johann Kranz. 2021. On the Effectiveness of Overt and Covert Interventions in Influencing Cookie Consent: Field Experimental Evidence, In Forty-Second International Conference on Information Systems. *ICIS 2021 Proceedings* 1, 1–17. https://aisel.aisnet.org/icis2021/cyber_security/cyber_security/5

[53] Giorgio Maone. 2023. NoScript. https://noscript.net/; Accessed on 2023.02.18.

[54] Arunesh Mathur, Jessica Vitak, Arvind Narayanan, and Marshini Chetty. 2018. Characterizing the Use of Browser-Based Blocking Extensions to Prevent Online Tracking. In *Proceedings of the Fourteenth USENIX Conference on Usable Privacy and Security (SOUPS '18)*. USENIX Association, USA, 103–116. event-place: Baltimore, MD, USA.

[55] C. Matte, N. Bielova, and C. Santos. 2020. Do Cookie Banners Respect my Choice? Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE Computer Society, Los Alamitos, CA, USA, 791–809. https://doi.org/10.1109/SP40000.2020.00076

[56] Philipp Mayring. 2014. *Qualitative content analysis: theoretical foundation, basic procedures and software solution.* AUT, Klagenfurt.

[57] Joseph E McGrath. 1981. Dilemmatics: The study of research choices and dilemmas. *American Behavioral Scientist* 25, 2 (1981), 179–210.

[58] Maryam Mehrnezhad, Kovila Coopamootoo, and Ehsan Toreini. 2022. How Can and Would People Protect From Online Tracking? *Proceedings on Privacy Enhancing Technologies* 2022, 1 (Jan. 2022), 105–125. https://doi.org/10.2478/popets-2022-0006

[59] Jakob Nielsen. 1992. Finding usability problems through heuristic evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, New York, NY, USA, 373–380.

[60] Jakob Nielsen and Rolf Molich. 1990. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, New York, NY, USA, 249–256.

[61] Midas Nouwens, Rolf Bagge, Janus Bager Kristensen, and Clemens Nylandsted Klokmose. 2022. Consent-O-Matic: Automatically Answering Consent Pop-Ups Using Adversarial Interoperability. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 238, 7 pages. https://doi.org/10.1145/3491101.3519683

[62] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. 2020. Dark Patterns after the GDPR: Scraping Consent Pop-Ups and Demonstrating Their Influence. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376321

[63] Maggie Oates, Yama Ahmadullah, Abigail Marsh, Chelse Swoopes, Shikun Zhang, Rebecca Balebako, and Lorrie Faith Cranor. 2018. Turtles, Locks, and Bathrooms: Understanding Mental Models of Privacy Through Illustration. *Proceedings on Privacy Enhancing Technologies* 2018, 4 (Oct. 2018), 5–32. https://doi.org/10.1515/popets-2018-0029

[64] Erica Olmsted-Hawala, Temika Holland, and Victor Quach. 2014. 3 - Usability Testing. In *Eye Tracking in User Experience Design*, Jennifer Romano Bergstrom and Andrew Jonathan Schall (Eds.). Morgan Kaufmann, Boston, 49–80. https://doi.org/10.1016/B978-0-12-408138-3.00003-0

[65] W. Palant. 2023. Adblock Plus. https://adblockplus.org/; Accessed on 2023.02.13.

[66] Gillian Porter, Tom Troscianko, and Iain D. Gilchrist. 2007. Effort during visual search and counting: Insights from pupillometry. *The Quarterly Journal of Experimental Psychology* 60 (2007), 211–229. https://doi.org/10.1080/17470210600673818

[67] Kopo M. Ramokapane, Anthony C. Mazeli, and Awais Rashid. 2019. Skip, Skip, Skip, Accept!!!: A Study on the Usability of Smartphone Manufacturer Provided Default Features and User Privacy. *Proceedings on Privacy Enhancing Technologies* 2019, 2 (April 2019), 209–227. https://doi.org/10.2478/popets-2019-0027

[68] Franziska Roesner, Tadayoshi Kohno, and David Wetherall. 2012. Detecting and Defending Against Third-Party Tracking on the Web. In *9th USENIX Symposium on Networked Systems Design and Implementation (NSDI 12)*. USENIX Association, San Jose, CA, 155–168. https://www.usenix.org/conference/nsdi12/technical-sessions/presentation/roesner

[69] Iskander Sanchez-Rola, Matteo Dell'Amico, Platon Kotzias, Davide Balzarotti, Leyla Bilge, Pierre-Antoine Vervier, and Igor Santos. 2019. "Can I Opt Out Yet?": GDPR and the Global Illusion of Cookie Control. In *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security* (Auckland, New Zealand) *(Asia CCS '19)*. Association for Computing Machinery, New York, NY, USA, 340–351. https://doi.org/10.1145/3321705.3329806

[70] Florian Schaub, Aditya Marella, Pranshu Kalvani, Blase Ur, Chao Pan, Emily Forney, and Lorrie Faith Cranor. 2016. Watching Them Watching Me: Browser Extensions Impact on User Privacy Awareness and Concern. In *Proceedings 2016 Workshop on Usable Security*. Internet Society, San Diego, CA, 1–10. https://doi.org/10.14722/usec.2016.23017

[71] D. Singer and R. Fielding. 2019. Tracking Preference Expression (DNT) W3C Working Group Note. https://www.w3.org/TR/tracking-dnt/.

[72] Michael Smith, Peter Snyder, Moritz Haller, Benjamin Livshits, Deian Stefan, and Hamed Haddadi. 2022. Blocked or Broken? Automatically Detecting When Privacy Interventions Break Websites. *Proceedings on Privacy Enhancing Technologies* 2022, 4 (Oct. 2022), 6–23. https://doi.org/10.56553/popets-2022-0096

[73] Than Htut Soe, Oda Elise Nordberg, Frode Guribye, and Marija Slavkovik. 2020. Circumvention by design – dark patterns in cookie consents for online news outlets. http://arxiv.org/abs/2006.13985 arXiv:2006.13985 [cs].

[74] Konstantinos Solomos, Panagiotis Ilia, Sotiris Ioannidis, and Nicolas Kourtellis. 2019. Clash of the Trackers: Measuring the Evolution of the Online Tracking Ecosystem. https://doi.org/10.48550/ARXIV.1907.12860

[75] StatCounter. 2023. Global market share held by leading internet browsers from January 2012 to January 2023. https://www.statista.com/statistics/268254/market-share-of-internet-browsers-worldwide-since-2009/

[76] Peter Story, Daniel Smullen, Yaxing Yao, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh, and Florian Schaub. 2021. Awareness, Adoption, and Misconceptions of Web Privacy Tools. *Proceedings on Privacy Enhancing Technologies* 2021, 3 (July 2021), 308–333. https://doi.org/10.2478/popets-2021-0049

[77] Alina Stöver, Nina Gerber, Christin Cornel, Mona Henz, Karola Marky, Verena Zimmermann, and Joachim Vogt. 2022. Website operators are not the enemy either - Analyzing options for creating cookie consent notices without dark patterns. In *Mensch und Computer 2022 - Workshopband*, Karola Marky, Uwe Grünefeld, and Thomas Kosch (Eds.). Gesellschaft für Informatik e.V., Bonn, 1–9. https://doi.org/10.18420/muc2022-mci-ws01-458

[78] Lisa M. Sullivan, Janice Weinberg, and John F. Keaney. 2016. Common Statistical Pitfalls in Basic Science Research. *Journal of the American Heart Association: Cardiovascular and Cerebrovascular Disease* 5, 10 (Sept. 2016), e004142. https://doi.org/10.1161/JAHA.116.004142

[79] David Thissen, Lynne Steinberg, and Daniel Kuang. 2002. Quick and Easy Implementation of the Benjamini-Hochberg Procedure for Controlling the False Positive Rate in Multiple Comparisons. *Journal of Educational and Behavioral Statistics* 27, 1 (March 2002), 77–83. https://doi.org/10.3102/10769986027001077

[80] Tobii. 2022. Tobii Pro Fusion. Retrieved 26 November 2022 from: https://www.tobii.com/products/eye-trackers/screen-based/tobii-pro-fusion.

[81] Tobii. 2022. Tobii Pro Lab - Software for experimental research. Retrieved 26 November 2022 from: https://www.tobii.com/products/software/data-analysis-tools/tobii-pro-lab.

[82] Michael Toth, Nataliia Bielova, and Vincent Roca. 2022. On dark patterns and manipulation of website publishers by CMPs. *Proceedings on Privacy Enhancing Technologies* 3 (2022), 478–497.

[83] Christine Utz, Sabrina Amft, Martin Degeling, Thorsten Holz, Sascha Fahl, and Florian Schaub. 2023. Privacy Rarely Considered: Exploring Considerations in the Adoption of Third-Party Services by Websites. *Proceedings on Privacy Enancing Technologies Symposium* 2023 (2023), 5–28. Issue 1. https://doi.org/10.56553/popets-2023-0002

[84] Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. 2019. (Un)informed Consent: Studying GDPR Consent Notices in the Field. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. ACM, London United Kingdom, 973–990. https://doi.org/10.1145/3319535.3354212

[85] Maggie Van Nortwick and Christo Wilson. 2022. Setting the Bar Low: Are Websites Complying With the Minimum Requirements of the CCPA? *Proceedings on Privacy Enhancing Technologies* 2022 (Jan. 2022), 608–628. https://doi.org/10.

2478/popets-2022-0030
[86] Colin Ware. 2021. Chapter Twelve - Designing Cognitively Efficient Visualizations. In *Information Visualization (Fourth Edition)*, Colin Ware (Ed.). Morgan Kaufmann, Burlington, Massachusetts, US, 425–456. https://doi.org/10.1016/B978-0-12-812875-6.00012-8
[87] Johannes Zagermann, Ulrike Pfeil, and Harald Reiterer. 2018. Studying Eye Movements as a Basis for Measuring Cognitive Load. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. ACM, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188628

A STUDY MATERIAL

Demographics questionnaire

- Age (collected in ranges)
- Gender (female, male, diverse, no response)
- Education (Finished School, Apprenticeship, High-school Diploma, University degree)
- Internet use for work (in hours)
- Internet use for free time (in hours)
- Experience with browser extensions (yes/no)
- Experience with privacy-preserving browser extensions (yes/no)
- List of used browser extensions (open text field)
- Affinity for Technology Interaction Scale (ATI, [29])
- Internet Users' Information Privacy Concerns Scale (IUIPC-8, [32])

Install questionnaire

- How participants think the browser extension works (open text field)
- What negative consequences using CookieBlock could bring with it (open text field)
- System Usability Scale (SUS, [14])
- Positive and negative experiences (open text field)

Questionnaire provided after each web shop task

- System Usability Scale (SUS, [14])
- Description of positive and negative experiences (open text field)

Final questionnaire

- How participants think the browser extension works (open text field)
- Whether people noticed CookieBlock caused a website breakage (yes/no/other)
- Satisfaction with extension overall (10-point Likert scale ranging from not at all satisfied to very satisfied)
- Intention to use when considering that websites work well in 85% of cases and that website breakage with impacted functionality concerns about 7% of cases (10-point Likert scale ranging from highly unlikely to highly likely)
- General CookieBlock suggestions (open text field)
- Suggestions for CookieBlock interface design improvement (open text field)
- Comments on an alternative interface design example shown in Fig. 4 (open text field)
- Additional comments (open text field)

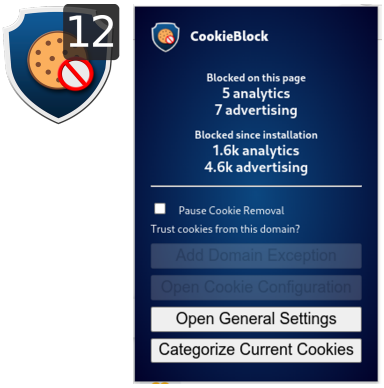


Figure 4: A screenshot of the suggested interface design change commented on by the participants.

B WEB SHOP SCREENSHOTS

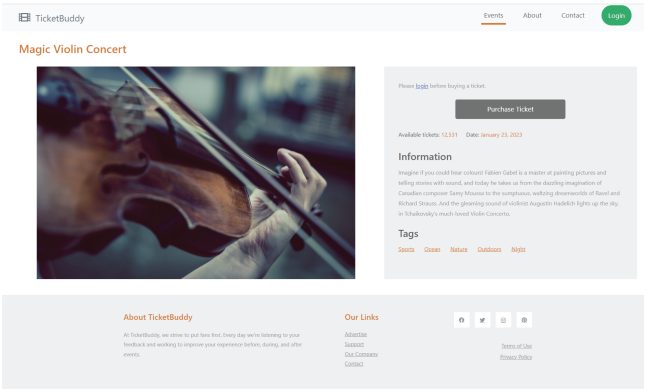


Figure 5: A screenshot of the "ticketbuddy" web shop.

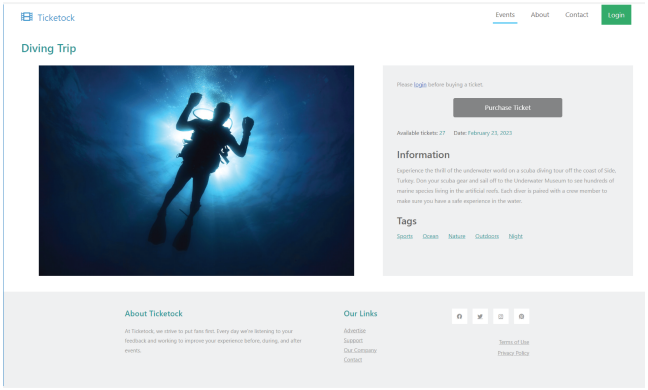


Figure 6: A screenshot of the "ticketock" web shop.

C OVERVIEW ON PRIVACY-RELATED BROWSER EXTENSIONS.

The following Table 4 provides an overview of different privacy-affecting browser extensions, highlighting the tracker blocking tools selected for the expert study. As a primary selection criterion from a user perspective, the anticipated privacy-usability trade-off is marked in bold font. The table furthermore shows differences in the functionalities and technologies of other tracker blockers as well as related privacy extensions with a different focus, such as ad blocking or the automatic handling of consent notices. The user numbers in the Chrome Web Store represent the current state on August 9, 2023.

Table 4: Comparison of CookieBlock with Expert-Selected Tracker-Blocking Tools, Including Previously Researched Privacy-Focused Tracker and Ad Blockers, and Related Consent Notice Extensions

Primary Purpose	**Tracker Blocking**					**Ad Blocking**		**Cookie Notice Handling**			
Tool	**CookieBlock [10]**	**Ghostery [24]**	**NoScript [53]**	**Privacy Badger [27]**	**DuckDuckGo Privacy Essentials [23]**	**uBlock Origin [22, 35]**	**AdBlock [1, 65]**	**Consent-O-Matic [41, 61]**	**Cookie Enforcer [43]**	**Cookie Glasses [55]**	**Cliqz Autoconsent [49, 50]**
Functionality	blocks only cookies	blocks third-party trackers and ads, and rejects cookies	blocks JavaScript, plugins, and other active content	blocks web requests	blocks trackers and rates websites	blocks ads and other content	blocks ads	automatic handling of consent notices	automatic discovery of cookie notices and decision on actions	compares registered consent to given consent	automatic opt-out of consent notices
Technology	machine learning	curated blocklists and heuristic elements	static algorithm based on user's exceptions	algorithm with periodic learning updates	curated crawling-based tracker blocklists	filter lists	filter lists	rule-based detection and enforcement	emulates user interaction with consent notices	compares to third-party databases	rule-based detection and enforcement
Disruptiveness	breakage of website functionality possible	breakage of website functionality very unlikely	breakage of website functionality likely	breakage of website functionality possible	only works with search engine DuckDuckGo, breakage of website functionality possible	breakage of website functionality rather unlikely	breakage of website functionality likely	requires specific CMPs to function correctly, but no impact on other website functionality	requires specific CMP configuration to function correctly, but no impact on other website functionality expected (not yet available)	information provision, no impact on website functionality anticipated	requires specific CMPs to function correctly, but no impact on other website functionality
Anticipated Privacy-Usability Trade-Off	**medium**	**low**	**high**	medium	medium	low	medium	low-medium	low-medium	low-medium	low-medium
Relevance/ Chrome Web Store Use	5.000+ users	2.000.000+ users	100.000+ users	1.000.000+ users	6.000.000+ users, only works with search engine DuckDuckGo	10.000.000+ users	10.000.000+ users	80.000+ users	not available on Chrome or for download, only available from ArXiv	254 users, project abandoned, technology outdated (only handles TCFv1)	4 users, not available in Chrome web store, only Firefox Experimental

D HEURISTIC EVALUATION

Table 5 provides a summary of the principles applied in the heuristic expert evaluation that were proposed by Nielsen [40]. Afterwards, Figure 7 provides screenshots of the three tools analysed in the heuristic evaluation, i.e., CookieBlock, NoScript, and Ghostery. The screenshots are supposed to facilitate understanding of the heuristic evaluation findings detailed for each principle in Tables 6, 7, and 8.

Table 5: Overview of the Usability Heuristics and Descriptions Used in Expert Evaluation

Usability Heuristic	Description
Visibility of system status	Ensure the extension provides clear and timely feedback on its current status, such as loading, processing, or error conditions. Users should always be aware of what the extension is doing.
Match between system and the real world	Use familiar and intuitive terminology, icons, and interactions that align with users' mental models and expectations. Avoid technical jargon or complex concepts that may confuse users.
User control and freedom	The extension should follow established browser conventions and standards, maintaining consistency in terms of interface elements, interactions, and terminology. Users should be able to transfer their knowledge from other extensions or browser features.
Consistency and standards	Use familiar and intuitive terminology, icons, and interactions that align with users' mental models and expectations. Avoid technical jargon or complex concepts that may confuse users.
Error prevention	Design the extension in a way that prevents errors or minimizes the likelihood of user mistakes. Provide clear error messages and offer options to undo or correct actions if possible.
Recognition rather than recall	Reduce the cognitive load on users by making information, options, and actions readily visible and accessible. Minimize the need for users to remember specific details or recall previous interactions by providing clear cues, labels, and context within the extension.
Flexibility and efficiency of use	Allow users to customize and personalize the extension to suit their needs. Provide shortcuts, customizable settings, and flexible workflows to accommodate different user preferences and skill levels.
Aesthetic and minimalist design	Strive for a clean and visually appealing interface. Avoid clutter, excessive elements, or distractions that could overwhelm or confuse users. Focus on essential features and prioritize information hierarchy.
Help users recognize, diagnose, and recover from website breakage	Design the extension to proactively prevent breakages when possible. If breakages do occur, provide informative and user-friendly error messages that explain the issue in plain language and suggest appropriate actions to resolve the problem. Offer clear guidance and recovery paths to help users get back on track.
Help and documentation	Include clear and accessible help documentation within the extension itself. Provide tooltips, contextual help, or a dedicated help section that explains the extension's features, functionality, and any potential limitations.

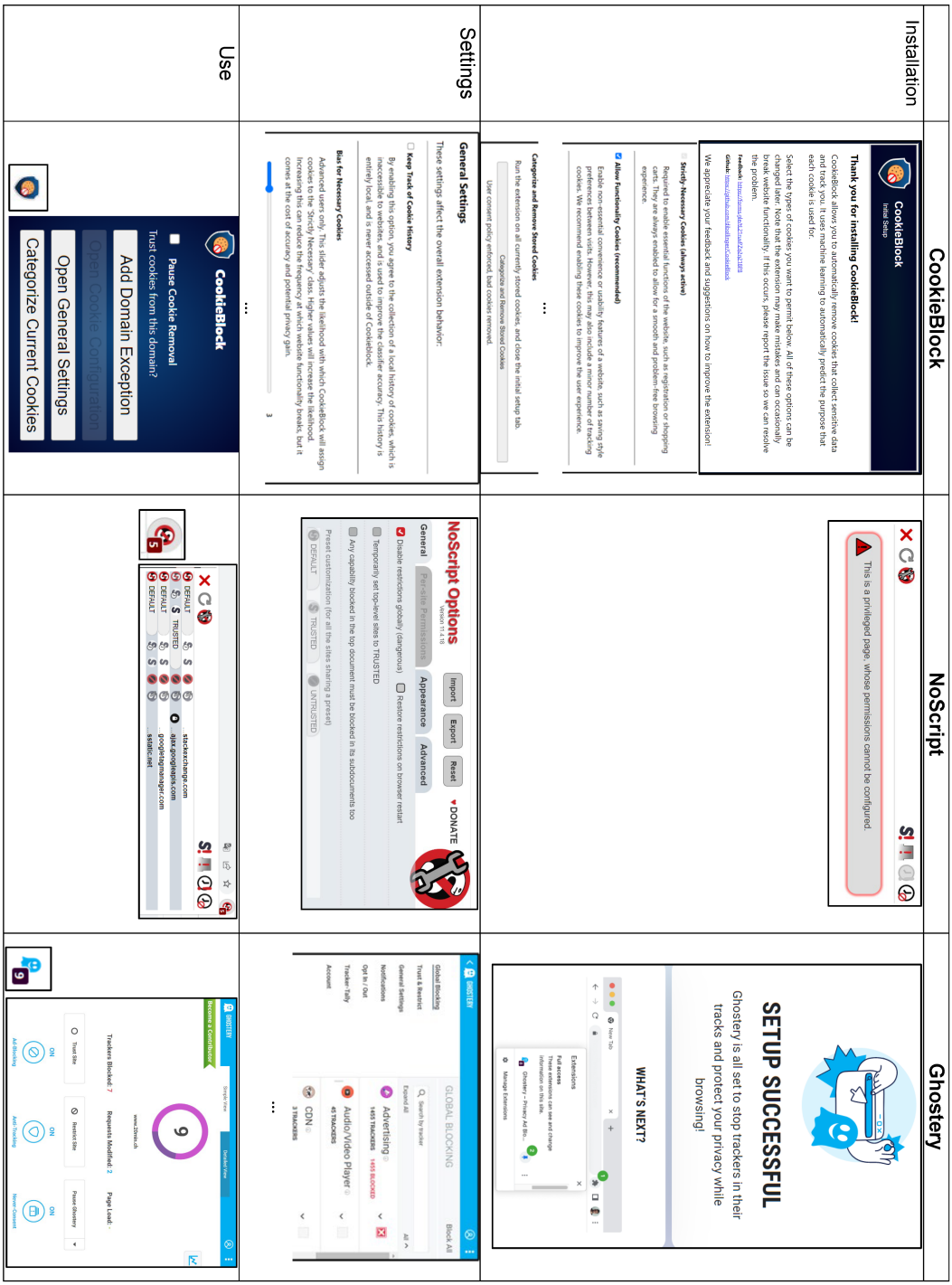


Figure 7: Illustrative screenshots of the installation, settings and use of the three tools CookieBlock, NoScript and Ghostery as analysed in the Heuristic Evaluation. Images from a) CookieBlock developed by [10] and available from https://github.com/dibollinger/CookieBlock, b) NoScript, Copyright held by Giorgio Maone, released under CC BY-SA 4.0 and available from https://noscript.net/, c) Ghostery, Copyright held by Ghostery GmbH and available from https://www.ghostery.com/. All images reproduced with permission.

Table 6: Summary of the findings of the heuristic evaluation separated for the tools CookieBlock, NoScript and Ghostery - Part 1

Usability Heuristic	CookieBlock	NoScript	Ghostery
Visibility of system status	**Installation**: Feedback on successful installation but not on whether Cookieblock is already active and pinned. Feedback on setting selection and removing cookies is difficult to note in normal font size, not really clear what happened. **Settings**: "Add Domain Exception" button switches to "Remove Domain Exception" when clicked and is greyed out when clicking not possible, greyed out button could be misunderstood as an error and difficult to understand why button is not always clickable -> indicate state via radio button? Difference between "Cookie Configuration" and "Settings" unclear, why is configuration not part of settings? Scale for Bias for Necessary Cookies not self-explainable. Unclear what happens when you choose "exempted". Sometimes not clear that something is clickable, e.g., known cookie list or cookie statistics **Use**: For lay users it might already be difficult to decide, in case of a problem, whether to pause cookie removal or to add domain exception. Feedback is only provided for pause cookie removal, then icon is greyed out. No feedback what happens on a specific website -> should at least be visible when clicking on icon.	**Installation**: Feedback on successful installation but not on whether NoScript is already active and pinned. **Settings**: Feedback that NoScript site cannot be configured because it's privileged helpful, but looks like warning message -> should look more neutral. Many different symbols of which meaning only becomes clearer through mouseover. Greyed-out symbols indicate non-clickable state, but then no mouseover available to understand meaning. Unclear what lock symbol means and what happens when it is clicked and turns red and unlocked, appears like an insecure state. Click on domain leads to error message rather than forwarding to domain. **Use**: Not clear what default setting is – trusted or untrusted? Feedback provided on which symbol is selected as it gets larger while the mouse hovers over one. Textual mouseover for each symbol available. Feedback on number of blocked things as compared to list that appears when clicking on symbol inconsistent, e.g., 6 things in list in pop-up and feedback that 7 blocked in symbol. Unclear why more blocked things appear when blocking temporarily deactivated. That symbol changes appearance when not active provides helpful feedback.	**Installation**: Start in inactive state and first needs to be enabled, but provides visual information how and where to pin as well as feedback that setup was successful. **Settings**: Visual feedback on selected settings, e.g., checkbox and red text next to selected type of blocked content **Use**: Feedback on number of trackers blocked in little number below symbol and in detailed screen when clicking on symbol. Actions are visible and feedback on implications of choices is provided. Visual color-feedback which choice is activated. Meaning of "Requests modified" unclear -> Is only ad blocked on website or other content, too?
Match between system and the real world	**Installation**: / **Settings**: General and advanced settings on one page -> should be better differentiated. Often technical language that might not be understandable, e.g., terms "classifier", "domain", "class". **Use**: /	**Installation**: / **Settings**: What is a privileged site? Naming and terms for options concerning default, trusted and untrusted is very technical and might be unclear to the user. **Use**: /	**Installation**: / **Settings**: / **Use**: The use of tracking categories like "Social Media" does not seem entirely clear and some trackers are simply "Unidentified." The never-consent option claims that it also blocks tracking, but in reality it only rejects choices in consent notice popups and does not directly interact with tracking.
User control and freedom	**Installation**: / **Settings**: When closing, the icon disappears. It first has to be pinned, which might be unclear to user. Accidentally added domains or typos can easily be removed by clicking on a cross, Info that bias configuration is only for expert users is quite hidden -> could be a separate area. **Use**: /	**Installation**: / **Settings**: If one clicks on close, icon disappears as it first has to be pinned, might be unclear to user. **Use**: Difficult to understand which functionality is available and what can be done in the settings menu.	**Installation**: / **Settings**: Information that Ghostery symbol first has to be pinned. Settings can be easily reversed, e.g., trust or block site. **Use**: /

Table 7: Summary of the findings of the heuristic evaluation separated for the tools CookieBlock, NoScript and Ghostery - Part 2

Usability Heuristic	CookieBlock	NoScript	Ghostery
Consistency and standards	**Installation**: Symbol on installation site and in browser consistent. The four cookie categories and the setting "keep track of cookie history" look very similar so that it looks as an additional category. Expectation that normally info on activating/pinning Cookieblock in installation process, but not given here. **Settings**: CookieBlock symbol is in expected place. **Use**: /	**Installation**: Expectation that normally info on activating/pinning NoScript in installation process, but not given here. Symbol of NoScript in installation process looks different from the symbol in browser. **Settings**: Symbols in pop-up are not intuitive (label in mouseover). No settings list or "general settings" as expected from platforms. **Use**: /	**Installation**: Symbol on installation site and in browser consistent. **Settings**: standards are considered, e.g. green and red color, underlined text for links and further information **Use**: In the settings menu, information that does not belong together is visually grouped together so that it appears to be belonging together.
Error prevention	**Installation**: While no errors seem apparent, unintended actions may be possible, e.g., by clicking on "Keep track of cookie history", because it looks like an additional cookie category. **Settings**: When click on cookie statistics, then all cookie data appears which looks like an error, no option to go back and no explanation but you can just close the tab. When you enter a domain exception you can remove it by clicking a cross which is an expected action. **Use**: No easily visible error recovery solution in case of website breakage as option for cookie pause is available but not obvious as solution for problem(connection between CookieBlock and website breakage not visible).	**Installation**: / **Settings**: Click on a domain leads to an error message, unclear whether to click on proceed or cancel. Implications of selections in advanced settings are unclear. **Use**: /	**Installation**: / **Settings**: / **Use**: If site is restricted, no sub settings or improvements can be made, site first has to be activated again, hierarchy of settings thus seems not plausible.
Recognition rather than recall	**Installation**: / **Settings**: Symbol for Cookieblock with shield can be recognized and does not need to be recalled. In settings screen lack of symbols and structure -> would be beneficial for recognition. **Use**: /	**Installation**: Symbols for NoScript are not consistent. **Settings**: Symbols are not recognizable without text. **Use**: Symbols are not recognizable but require further explanation.	**Installation**: / **Settings**: Symbols are consistently used and explained (and info buttons available in many places). **Use**: /

Table 8: Summary of the findings of the heuristic evaluation separated for the tools CookieBlock, NoScript and Ghostery - Part 3

Usability Heuristic	CookieBlock	NoScript	Ghostery
Flexibility and efficiency of use	**Installation**: / **Settings**: Configuration of cookies very detailed and complex for lay users -> better differentiation for lay and expert user would help to increase flexibility, all expert settings could be hidden or below lay user setting. **Use**: /	**Installation**: / **Settings**: Settings use tabs for easy overview. Advanced settings are separated from general settings. Search function allows for shortcuts for experienced users. **Use**: /	**Installation**: / **Settings**: / **Use**: : It seems easy to block or trust overall, but difficult to set granular advanced settings. "Submit a tracker" seems to be an advanced feature, so unclear why positioned in main menu.
Aesthetic and minimalist design	**Installation**: Information on how to provide feedback and suggestions on CookieBlock before actual cookie selection even though the latter is probably the user's priority. **Settings**: Settings and pop-up when clicking on icon is minimalist -> Setting screen could profit from structure, e.g., lay user and expert settings. Design of "cookie configuration" screen is not minimalist but has very detailed settings. **Use**: /	**Installation**: / **Settings**: While minimalist in the sense that there is little text, there are many symbols that are not self-explanatory. Setting to be able to choose dark mode good in terms of accessibility. Change of symbol might not be necessary and is not consistently applied across screens. **Use**: /	**Installation**: Visual illustration of installation process. Aesthetic and modern design. Consistent use of symbols. **Settings**: List and structure so that not all information is displayed at once. **Use**: /
Help users recognize, diagnose, and recover from errors	**Installation**: / **Settings**: / **Use**: Not clear that website breakage is related to CookieBlock from interface or feedback, no easily visible way out.	**Installation**: / **Settings**: / **Use**: When content is blocked, NoScript symbol appears on blocked content (e.g., video), so connection between blocking and breakage can be made. However, no information what to do or how to recover.	**Installation**: / **Settings**: / **Use**: Perhaps not clear that there is a connection between Ghostery's blocking functionality and potential website breakage, but Ghostery does not seem disruptive in comparison.
Help and documentation	**Installation**: Cookie types are explained on initial set-up page, but more info or different wording could be helpful in some cases, e.g., button "Categorize and remove stored cookies" . **Settings**: Information text available for each setting, but some explanations require expert knowledge. **Use**: No information on what CookieBlock is currently doing on website.	**Installation**:/ **Settings**: No further help or information available on first glance. **Use**: /	**Installation**: Help module and information on why and how to pin extension available. **Settings**: / **Use**: In many places little information symbols and further information available.

E ADDITIONAL DEMOGRAPHIC METRICS, EYE TRACKING, AND STATISTICS TABLES

Table 9: Means (SDs) of the IUIPC-10 Sub-Scales

Subscale	User Study Sample	"Sample V" from [32]
Control	5.84 (0.87)	5.86 (0.84)
Awareness	6.46 (0.75)	6.43 (0.66)
Collection	5.45 (1.04)	5.60 (1.04)

Table 10: Eye Tracking Measures

Measure	Purpose
Heat map	Focus comparisons between participants
Fixation time	How much processing an element invokes
First fixation	How visually apparent elements are
First click	How obvious the element's purpose is
Saccade number	How high mental workload is
Saccade velocity	How high mental workload is
Pupil diameter	How high mental workload is

Table 11: Detailed Eye Tracking Results

Measure	Icon	Dropdown	Significance
Fixation time	4.36 seconds	0.49 seconds	$p < .001$
Time to first fixation	2.71 minutes	1.96 minutes	$p = .712$
Time to first fixation (no help)	1.15 minutes	1.16 minutes	$p = .955$
Time to first fixation (help)	4.62 minutes	4.90 minutes	$p = .277$
Measure	Baseline	Breaking Task	Significance
Number of saccades	1.388 per second	2.294 per second	$p = .002$
Pupil diameter	3.20 mm	3.51 mm	$p < .001$
Measure	Icon Entry	Drop-down Exit	Significance
Saccadic Peak Velocity	132.28°/s	252.96°/s	$p = .012$

Table 12: Test Statistic Results

Research Question	Measure	Comparison	Test	Result
RQ2: Mental Model	Mental model accuracy	Required help	Fisher's	$p = .127$
		Breaking task time	ANOVA	$F(2, 37) = 1.14, p = .331$
	Prediction accuracy	Required help	Fisher's	$p = .046$
		Breaking task time	ANOVA	$F(2, 35) = 3.32, p = .096$
RQ3: User Interaction	Require Help	Task Speed	T Test	$t(33.65) = 5.19, p < .001, d = -1.66$
		Age	Fisher's	$p = .076$
		ATI	T Test	$t(36.10) = 0.19, p = .853, d = .06$
RQ4:User Evaluation	Breaking Condition SUS	Required Help	Wilcoxon	$z = 233, p = .220$
		Working Condition SUS	T Test	$t(74.88) = -3.06, p = .003, d = -.69$
		Install SUS	T Test	$t(71.19) = -3.56, p < .001, d = -.80$