Evaluating the Usability of Privacy Choice Mechanisms

Hana Habib and Lorrie Faith Cranor, *Carnegie Mellon University*

https://www.usenix.org/conference/soups2022/presentation/habib

This paper is included in the Proceedings of the Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022).

August 8–9, 2022 • Boston, MA, USA

978-1-939133-30-4

Open access to the Proceedings of the Eighteenth Symposium on Usable Privacy and Security is sponsored by USENIX.

Evaluating the Usability of Privacy Choice Mechanisms

Hana Habib
Carnegie Mellon University

Lorrie Faith Cranor
Carnegie Mellon University

Abstract

Privacy choice interfaces commonly take the form of cookie consent banners, advertising choices, sharing settings, and prompts to enable location and other system services. However, a growing body of research has repeatedly demonstrated that existing consent and privacy choice mechanisms are difficult for people to use. Our work synthesizes the approaches used in prior usability evaluations of privacy choice interactions and contributes a framework for conducting future evaluations. We first identify a comprehensive definition of usability for the privacy-choice context consisting of seven aspects: user needs, ability & effort, awareness, comprehension, sentiment, decision reversal, and nudging patterns. We then classify research methods and study designs for performing privacy choice usability evaluations. Next, we draw on classic approaches to usability testing and prior work in this space to identify a framework that can be applied to evaluations of different types of privacy choice interactions. Usability evaluations applying this framework can yield design recommendations that would improve the usability of these choice mechanisms, ameliorating some of the considerable user burden involved in privacy management.

1 Introduction

Consumer privacy protection has long been rooted in the notice and choice paradigm. This model assumes that companies notify users about how they handle their data and consumers exercise privacy choices according to their preferences. Thus, companies implement web and app interfaces with privacy choice mechanisms that allow users to make choices about some form of collection or use of their personal data, including device permission prompts, cookie consent notices, social media audience settings, targeted advertising opt-outs, and mailing list opt-outs. The possible design space of privacy choice mechanisms is broad, resulting in interfaces that vary in type of choice, functionality, timing, channel, and modality [15]. Despite the availability of privacy controls, the notice and choice model arguably has not resulted in effective consumer privacy protection, in part due to the poor usability of privacy choice mechanisms [53].

The design of privacy choice and consent interfaces can significantly impact users' privacy outcomes. Historically, companies have had economic motivation to encourage users to share their data through such interactions and may not have exerted more than minimal effort in testing the usability of their privacy choice and consent interfaces. Furthermore, privacy choice interfaces require usability considerations beyond those considered for typical user interfaces. Generally, users make privacy decisions when trying to accomplish a different goal (e.g., browse a website or make an online purchase), which means that a choice interface that interferes with the primary goal might score high with respect to the usability of the privacy decision but low with respect to the primary goal.

Prior usability evaluations of privacy choice mechanisms have highlighted several obstacles to their effective use. For example, some privacy choice mechanisms may be difficult to configure without substantial technical knowledge [36]. Some seem to require that users put aside their preconceived assumptions and read explanations that most users readily skip over [51]. Furthermore, the use of dark patterns may nudge users toward less privacy-protective options provided in the interface [56]. Prior studies often include actionable design recommendations for a particular privacy choice context (e.g., [21, 40, 60]).

The expanding literature on privacy choice interfaces has explored a variety of usability considerations for privacy choice interactions, utilizing a spectrum of usability testing methods from the field of human-computer interaction. In this

USENIX Symposium on Usable Privacy and Security (SOUPS) 2022.
August 7–9, 2022, Boston, MA, United States.

work, we distill the usability aspects explored and methods used in prior work into a framework that can inform the design of future usability evaluations of privacy choice interactions. To develop this framework we adopt our prior work [20] presenting a comprehensive definition of usability for the context of privacy choice mechanisms consisting of seven objectives: user needs, ability & effort, awareness, comprehension, sentiment, decision reversal, and nudging patterns. We then categorize different research methods and study designs that can be used to perform usability evaluations of privacy choice interfaces. Next, drawing on classic approaches to usability testing and prior evaluations of privacy controls, we construct the Privacy Choice Evaluation Framework that can be applied to future evaluations of privacy choice interfaces. The framework provides criteria for evaluating each aspect of usability through the relevant evaluation approaches, serving as a guide for organizations that want to ensure provided privacy controls are effective in enabling consumers to manage their privacy. Furthermore, regulators can make use of this framework as they work to hold companies accountable to rigorous usability testing of privacy choice and consent processes.

After presenting our framework, we present an overview of the literature on privacy choice evaluations, illustrating the applicability of our framework. We then discuss additional considerations, guidance for organizations, and limitations of privacy choice usability. Our appendix includes guidance on using the evaluation framework through a detailed example.

2 Defining Privacy Choice Usability

To consider the holistic usability of privacy choice interfaces, it is important to first identify aspects of usability that are relevant to the privacy choice experience. We adopt our previous work [20], which reviewed definitions of usability drawn from academics and practitioners in the privacy, HCI, and user experience (UX) fields and identified seven distinct aspects of privacy choice usability. We use these seven aspects of usability to provide an organizing structure for our framework.

User Needs: Whether a privacy choice interface addresses the intended users' privacy needs in a particular privacy choice context. Also includes accuracy and completeness of the interface in addressing these needs. *Components from previous definitions:* **Effectiveness** (Feng et al. [15], ISO [27], Quesenbery [52]), **Useful** (Schaub and Cranor [54], Morville UX Honeycomb [44])

User Ability & Effort: Whether a privacy choice interface allows the intended users to accomplish a particular privacy goal and with minimal effort. *Components from previous definitions:* **Efficiency** (Feng et al. [15], ISO [27], Quesenbery [52], Nielsen [45]), **Usable** (Schaub and Cranor [54], Morville UX Honeycomb [44]), **Accessible** by "non-experts" (Morville UX Honeycomb [44])

User Awareness: Whether the intended users are aware that a particular privacy choice exists within a privacy choice interface, and if they are able to find it. *Components from previous definitions:* **User awareness** (Feng et al. [15]), **Findable** (Schaub and Cranor [54], Morville UX Honeycomb [44]), **Easy to learn** - initial orientation (Quesenbery [52], Nielsen [45])

User Comprehension: Whether the intended users understand what a particular privacy choice does and the implications of their decisions. *Components from previous definitions:* **Comprehensiveness** (Feng et al. [15]), **Understandability** (Schaub and Cranor [54]), **Easy to learn** - continued learning (Quesenbery [52])

User Sentiment: Whether the intended users are satisfied with a privacy choice interface and options it provides. This includes whether users have faith that the privacy choice will be honored. *Components from previous definitions:* **Satisfaction** (ISO [27], Nielsen [45]), **Engaging** (Quesenbery [52]), **Desirable** (Morville UX Honeycomb [44]), **Credible** (Morville UX Honeycomb [44])

Decision Reversal: Whether a privacy choice interface allows the intended users to correct an error or change their decision. This also includes the effort required to do so. *Components from previous definitions:* **Error tolerant** (Quesenbery [52], Nielsen [45])

Nudging Patterns: Whether the design of a privacy choice interface leads the intended users to select certain choices in the interface over others (including dark patterns that lead users to less privacy-protective options). *Components from previous definitions:* **Neutrality** (Feng et al. [15])

3 Privacy Choice Evaluation Approaches

This section describes research methods and study designs that can be applied to privacy choice evaluations. While it is not a comprehensive list of all possible evaluation techniques, it demonstrates a wide breadth and diversity of approaches.

3.1 Expert Evaluation Methods

Inspection-based approaches, in which usability obstacles are identified through a systematic review of the interface by a domain expert, can be adapted to evaluate the usability of privacy choice and consent interfaces. Such approaches may be particularly beneficial in evaluating privacy choice interfaces in contexts where users may lack requisite background

privacy knowledge or experience. Prior examples of privacy choice usability studies conducted through expert evaluation include Grey et al.'s interaction criticism approach and Soe et al.'s heuristic evaluation of cookie consent banners [19, 56]. Here we provide a brief description of five inspection-based methods that could be used in evaluating for different usability aspects. Additional information about these approaches can be found in the HCI literature (e.g., [66]).

Perspective-based UI Inspection: One or more people evaluate the privacy choice interface from the perspective of a particular type of user (super-user, less-tech savvy, person with disability) or through the lens of a specific normative value, in this case privacy.

Individual Expert Review: One or more experts in HCI, the privacy choice domain, or the product conducts a review to find usability problems in a privacy choice interface according to the usability aspect(s) being evaluated.

Cognitive Walkthrough: An expert or team interacts with a privacy choice interface to identify usability issues that primarily impact user awareness. This method is based on the theory that users learn through exploration.

Heuristic Evaluation: An individual or team evaluates a privacy choice interface design against a list of UX principles (e.g. Nielsen Heuristics [46]) or other pre-defined criteria (e.g., regulatory requirements).

Formal Usability Evaluation: Trained inspectors conduct coordinated, individual usability assessments of a privacy choice interface (similar to formal code inspections). This may include collecting information about the shortest path, minimum number of actions, and time taken to complete a privacy choice task.

3.2 User Study Designs

User studies provide perspectives from individuals who are more likely to represent the opinions and behaviors of end-users of the privacy choice interface. Such evaluations of privacy choice interfaces can be implemented through different research methods and study designs as outlined below. Studies may combine elements to explore how well a privacy choice interface addresses particular usability aspects.

3.2.1 No Task Assigned

Self-reported: Self-report methods can help with understanding users' experiences with a privacy choice interface in the context of their actual use of the system. This can provide valuable insight even for privacy interfaces that users may encounter infrequently. Furthermore, self-report methods can help understand users' privacy needs for a particular context. These studies can be conducted through surveys, interviews, and focus groups utilizing qualitative prompts, measurement scales, and other question types. Examples of prior self-report studies related to privacy choice interfaces include Malkin et al.'s survey of smart speaker users [40] and Colnago et al.'s interview study informing the design of a privacy assistant [11].

Observed: Observation studies primarily involve measurement of users' behavior when interacting with a deployed privacy choice interface, sometimes as part of an A/B test. Examples of such metrics include the average amount of time spent before making a privacy choice or percentage of users who click a particular option. Such studies provide an advantage over other study designs by providing insight into when and how users are actually interacting with an interface, which is particularly useful for the privacy choice context as privacy management is typically not users' primary reason for engaging with a system. However, observation studies do not typically provide an explanation as to why users interact with it in the way that they do, unless paired with an interview or survey. Previous observation studies of cookie consent interface designs include Utz et al.'s field study evaluation [61] and the logistics company DHL's A/B tests [49].

3.2.2 Participants Assigned Privacy Task

In their natural use of a system, users may encounter a particular privacy choice interface so infrequently that it may be difficult for researchers to assess its usability. Thus evaluating for some usability aspects may require explicitly assigning privacy-related tasks to ensure that users interact with the interface being evaluated. Additionally, participants are typically asked questions before or after task completion (or both). These user studies can be implemented through surveys, experiments, or lab usability studies.

Hypothetical Privacy Scenario: Participants are given a realistic scenario motivating a privacy choice and are asked how they would use a privacy choice interface (or other mechanism) to make that choice. An example of a hypothetical scenario that was used by Habib et al. [21] and Kumar et al. [5] to introduce tasks involving email opt-outs is "You just got the 10th update email from this website today. Now you want to stop receiving them."

Participant Inspection: Participants are shown a privacy choice interface and are encouraged to fully engage with it prior to answering questions (e.g., to measure their awareness or comprehension). Typically, participants are allowed to reference the interface while they are answering questions. Tsai

et al.'s online experiment used participant inspection to compare the design of a new Android permission manager tool with Android's native permission management interface [60].

Participant Quick Review: Participants are shown a privacy choice interface but are only allowed a short period to engage with it (e.g., 3 seconds). Typically, participants are not allowed to reference the interface while they are answering questions. Quick review may also be done as part of a task in which participants are exposed to the interface, but answer questions about it after they complete the task and the interface is no longer in view. Cranor et al. used quick review to evaluate whether participants noticed a "Do Not Sell My Personal Information" opt-out link and icon in the footer of an e-commerce website after their attention was directed to a nearby link [12].

Make Personal Privacy Choices: Participants are shown a privacy choice interface and are asked how they would interact with it according to their own personal privacy preferences. For example, Krsek et al.'s experiment asked participants to select their preferences for Facebook privacy settings under different nudging conditions [33].

3.2.3 Participants Assigned Distraction Task

Considering that privacy/security are often secondary priorities when users interact with a system, simulating this in an online experiment or lab usability study might require assigning participants a "distraction task." Examples of distraction tasks include shopping for a particular item, or finding information on a website. Participants should encounter the choice interface or an indicator leading to it during their task.

Privacy Choice Prompt Appears: Participants are asked to complete a task that is unrelated to the privacy choice interface being evaluated, but are exposed to the privacy choice interface at some point in the study. For example, in Bermejo Fernández et al.'s online experiment evaluating the usability of cookie consent interfaces, a cookie consent banner appeared as participants arrived at the website to complete a survey about smart home devices [6].

Participant Seeks Out Privacy Settings: Participants are asked to complete a task that is unrelated to privacy but as part of the interface they can see the current privacy settings. During the course of task completion they may choose to change their privacy settings according to their preferences. Vaniea et al. conducted a series of lab studies in which participants were assigned photo management tasks during which they had the opportunity to observe and change the access control settings for each photo [62]. However, the authors report a number of challenges they encountered while conducting these studies using this approach, including making sure participants understood the somewhat-complex desired access control policy, and balancing the need to make participants aware of the access control settings with a desire not to prime participants to think about access control more than they normally would.

4 The Privacy Choice Evaluation Framework

We introduce the Privacy Choice Evaluation Framework, summarized in Table 1, which provides a set of criteria that can be used in usability evaluations of privacy choice interfaces. We structure the framework according to the seven usability aspects defined in Section 2. For each criterion included in the framework, we highlight the study approaches described in Section 3 and describe measures or example prompts that can be incorporated into a usability study. We refer to established usability metrics and heuristics when appropriate, or specific components of existing usability scales that are applicable to the privacy choice context. It is important to note that the usability requirements and acceptable thresholds for meeting them are not universal, but rather depend on the context of the privacy choice interface. Many factors, including intended user groups, complexity of options, and devices used to display the privacy choice interface, influence whether a given privacy choice interface is sufficiently usable. The framework also considers the types of privacy choice interfaces relevant to each criterion, in terms of the Timing component of the privacy choices design space: *on-demand* (privacy settings pages that the user seeks out) or *interruptive* (privacy choice interfaces that appear at setup, just-in-time, are context-aware, are periodic, or are personalized) [15]. Furthermore, we provide citations to prior privacy choice evaluations when applicable to demonstrate possible implementations of the listed criteria.

4.1 User Needs

Prior to designing an interface, design teams often complete a needs assessment using qualitative approaches to better understand how and why users might use the interface. It is important to assess whether a resulting interface design is aligned with the identified needs and how completely it addresses them. Assessing user needs is relevant to both interruptive and on-demand privacy choice interfaces. Some evaluations in other parts of this framework rely on an understanding of user needs associated with a privacy interface.

4.1.1 Users' Privacy Objectives

This criterion pertains to understanding users' privacy objectives when using a particular system. Assessments of users' privacy objectives can be conducted as self-reported evaluations of past experiences or user studies involving assigned tasks. Prior work evaluating users' privacy objective when

Framework Criterion	Usability Aspect							Evaluation Approach					Interface Timing	
	Needs	Ability & Effort	Awareness	Comprehension	Sentiment	Decision Reversal	Nudging Patterns	Expert Evaluation	Self-Report	Observation	Privacy Task	Distraction Task	Interruptive	On-demand
Users' privacy objectives	●						●		●		●	●	●	●
Users' intentions	●								●		●	●	●	●
Interface completeness	●							●					●	●
Interface accuracy	●							●					●	●
Ability - make privacy choice		●				●	●			●	●	●	●	●
Time taken - make privacy choice		●				●	●			●	●	●	●	●
User actions - make privacy choice		●				●	●			●	●	●	●	●
Perceived effort - make privacy choice		●				●			●		●	●	●	●
Estimated effort - make privacy choice		●				●		●					●	●
Awareness of choice existence			●			●	●		●			●	●	●
Ability - find privacy choice			●			●				●	●			●
Time taken - find privacy choice			●			●				●	●			●
User actions - find privacy choice			●			●				●	●			●
Perceived effort - find privacy choice			●			●			●		●			●
Estimated effort - find privacy choice			●			●		●						●
Objective knowledge - focused attention				●		●	●		●		●		●	●
Objective knowledge - unfocused attention				●			●		●			●	●	
Perceived effort - comprehension				●		●			●		●	●	●	●
Estimated effort - comprehension				●		●		●					●	●
Perceived transparency & control					●		●		●		●	●	●	●
Subjective knowledge					●		●		●		●	●	●	●
Levels of comfort & trust					●		●		●		●	●	●	●
Investment in decision-making					●		●		●		●	●	●	●
Impact on individual welfare							●		●		●	●	●	●
Unintended societal consequences							●	●					●	●
Alignment with regulatory objectives							●	●					●	●
Individual autonomy							●		●	●	●	●	●	●

Table 1: A summary of the Privacy Choice Evaluation Framework which provides an overview of the evaluation critera (grouped by the usability aspect in Section 4 under which they are described). Marked are the applicable usability aspects (defined in Section 2), evaluation approaches (described in Section 3), and timing of privacy choice interface (interruptive and/or on-demand) for each criterion.

using a privacy choice interface includes Fiesler et al.'s survey of Facebook users, which asked "Why did you choose this privacy setting?" for each post shared by their participants. Additional example prompts include:

- What settings or controls related to [domain of privacy choice] would you like to have available to you, if any? [for initial exploration into user needs prior to designing the privacy choice interface]
- What *other* settings or controls related to [domain of privacy choice] would you like to have available to you, if any? [for further exploration into user needs related to an existing privacy choice interface design]

4.1.2 Users' Intentions

Similar to exploring users' objectives, it is important to assess why users interact with privacy choice interfaces in the way that they do, including evaluating users' decision strategies. Assessing users' intentions requires participants to reflect on what they were trying to achieve in a past interaction with a privacy choice interface, which could be privacy related (e.g., trying to prevent a certain type of data collection) or more practical (e.g., to continue to the main website). This can be conducted as self-reported evaluation of past experiences or user studies involving assigned tasks. An example prompt to assess users' intentions is: What were you trying to achieve when you [interacted with the choice interface]?

4.1.3 Interface Completeness

The criterion assesses how completely an implemented privacy choice interface achieves users' needs through an expert evaluation. This requires having some knowledge of users' objectives through a user study and thus ideally should be done in conjunction with the criterion described in 4.1.1. Such evaluations could include heuristics such as:

- Does the interface meet the needs of different types of users (e.g. those who want fine-grained controls and those who want simplicity.)?
- Does it allow users to achieve all of their stated objectives, or only some of them?

Some interfaces may be incomplete because they do not allow users to make desired privacy choices at all, for example not offering the option to post anonymously on a social media platform. Others may offer desired choices, but not at the level of granularity desired by some users, for example allowing social media users to restrict the audience of their posts to friends, but not allowing them to restrict the audience to only a particular subset of their friends.

4.1.4 Interface Accuracy

In addition to how completely a privacy choice interface meets users' needs, an expert evaluation can also assess how accurately it achieves users' needs. This requires having some knowledge of users' intentions when using a privacy choice interface, and could be done in conjunction with a user study exploring the criterion described in 4.1.2. These evaluations could include evaluating whether there is a mismatch between what the user said they were trying to achieve and what the interface actually does, and identifying how the interface helps users accomplish their goals. For example, some cookie consent interfaces give users a choice of "accept all cookies," "reject all cookies," or "manage cookies." Reject all cookies is not an accurate label on most websites where it actually rejects all non-essential cookies but not the "strictly necessary" cookies needed for the site to function and which are permitted under GDPR.

4.2 User Ability & Effort

Usability testing often involves quantitative measures that estimate the effort involved in using an interface. These metrics can be used to compare interfaces (e.g., a previous version of the interface, alternate designs, or the interface of a similar product). Measuring perceived effort is relevant to both interruptive and on-demand choice interfaces. For on-demand privacy choice interfaces, much of the effort involved in using the interface will likely be in finding where it is (which we discuss as separate criteria in 4.3), but users could possibly make other errors such as forgetting to save their choices or toggling a choice in the wrong direction.

4.2.1 Ability to Make a Privacy Choice

Ability evaluations may assess whether users are able to complete the end-to-end interaction required to make a privacy choice, as well as the type and extent of assistance they require. Ability to make a privacy choice can be measured through observational field studies or user studies involving task assignment. Prior work evaluating for ability to make a privacy choice includes Chalhoub et al.'s ethnographic study which surfaced participants' inability to configure privacy settings on their smart home devices [7].

4.2.2 Time Taken to Make Privacy Choice

Time is one measure of the effort required to use an interface, and can be measured through both observation and user study tasks. However, the raw time to complete a privacy decision may be an imperfect measure if users are multitasking or thinking aloud during a moderated study. Alternative time-based metrics include time-based efficiency and overall relative efficiency [43]. An example of prior work

that included timing metrics in their usability testing is Garlach and Suthers's study evaluating the effectiveness of the AdChoices icon in the mobile environment [18].

4.2.3 User Actions Required to Make Privacy Choice

Another measure of effort is the number and type of user actions (e.g., clicks, hovers, form fields) required to complete a privacy choice. This may sometimes be a more reliable measure than time and may also reveal common user errors that result in extra user actions. User actions can be measured through observation as well as user study tasks. Habib et al. tracked clicks, scrolls, form field, check boxes, and hovers in a lab usability study of opt-out and data deletion interfaces [21].

4.2.4 Perceived Effort in Making a Privacy Choice

After completing a task that requires using a privacy choice interface, participants can be asked questions related to the perceived ease or difficulty of their experience. Alternatively, these questions can be asked about participants' prior experiences with a privacy choice interface outside of the study environment. Work by Tsai et al. and Habib et al. reported perceived effort by asking participants a version of the the Single Ease Question (SEQ) ("Overall, how easy or difficult was it to perform this task?") to evaluate different privacy choice interfaces [21, 22, 60]. Other commonly used prompts that measure perceived effort on a Likert scale include items 2, 3, 4, and 8 on the System Usability Scale (SUS) [35].

4.2.5 Estimated Effort Required to Make a Choice

Expert evaluation approaches can be used to estimate users' ability and effort in using a privacy choice interface to accomplish a particular goal. Such evaluations may include a set of design heuristics specific to the privacy choice interface or established usability heuristics (e.g., items 1-3, 7, 8 of the Nielsen heuristics [46]). Estimating ability and effort could also be done in conjunction with 4.2.2 and 4.2.3, as it may be helpful to compare the ability and effort of an "expert" with prior knowledge of the privacy choice interaction to those of user study participants. Habib et al. estimated the effort involved in using privacy opt-outs and data deletion mechanisms by counting the user actions in the shortest interaction path required to opt-out or delete data [24].

4.3 User Awareness

For privacy choice interfaces to be usable, it is necessary to ensure that users recognize that the privacy choice(s) exist and that they are able to find them. Awareness may be measured together or separately from user ability & effort (Section 4.2) as it is part of the interaction required to use a privacy choice interface. Testing for awareness may be less important for interfaces that interrupt the user's primary goal, compared to on-demand privacy settings pages that users must seek out. Furthermore, for step-wise privacy choice interfaces, in which choices are incrementally revealed, it may be sufficient to evaluate whether users are aware of the general types of options available, rather than every option offered in the interface.

4.3.1 Awareness of Choice Existence

Assessing awareness of privacy choice interfaces and available options, sometimes referred to as *discoverability* [3], requires study participants to have prior experience with the system but not necessarily the particular interface being evaluated. Thus, self-report evaluations or user studies with distraction tasks are appropriate for evaluating awareness. For interruptive interfaces, evaluating for this criterion might include whether participants can recall the specific choice interface or available privacy options, whether participants realized they were asked to make a privacy choice during a distraction task, and if can they identify which choice they made. For on-demand privacy choices, users might be asked about their own privacy objectives or told about objectives that some users have, and then asked whether they think there is an interface that might help them achieve this objective (as a follow-up researchers may then assess the users' ability to find it). An example of prior work measuring awareness is Cranor et al.'s study that evaluated whether participants noticed an opt-out link and icon present on the page [12].

4.3.2 Ability to Find Privacy Choice

This criterion can be incorporated into user studies that implement the criterion described in 4.2.1, as finding the privacy choice interface is typically the bulk of a privacy choice interaction for on-demand privacy choices. It may include assessing whether participants were able to find the choice interface without assistance, and for moderated studies, what hints aided participants in finding the privacy choice.

4.3.3 Time Taken to Find the Privacy Choice

Similarly, this criterion can be studied with the criterion described in 4.2.2. For example, Garlach and Suthers report the time taken by their study participants to find the AdChoices icon on a mobile device [18].

4.3.4 User Actions Taken to Find the Privacy Choice

Participants' interaction path while trying to find the privacy choice can also be studied alongside the criterion in 4.2.3.

4.3.5 Perceived Effort in Finding the Privacy Choice

This criterion is similar to that described in 4.2.4. After completing a study task that requires participants to seek out the privacy choice interface, participants can be asked questions

related to the perceived ease or difficulty in finding the privacy choice. For example, participants in Chen et al.'s study were asked to rate the difficulty of finding different app privacy settings [8]. Alternatively, participants can be asked about prior experiences with a privacy choice interface outside of the study environment in self-report studies.

4.3.6 Estimated Effort in Finding the Privacy Choice

Expert evaluation approaches can be used to estimate the difficulty of finding a privacy choice. Cognitive walkthroughs of the system may be especially relevant when evaluating the learnability of the privacy choice interaction [66]. Established usability heuristics (e.g., items 4 and 6 of the Nielsen heuristics [46]) also address findability. Estimating effort in finding the privacy choice could be done in conjunction with 4.3.3 and 4.3.4. Similar to the criterion described in 4.2.5, comparing the ability of an "expert" with prior knowledge of the system with those of study participants to find the privacy choice interface may suggest usability issues in the interaction if there is a large gap.

4.4 User Comprehension

For a privacy choice interface to be effective, it is important to ensure that users understand what it does and identify any misconceptions. When evaluating for comprehension, it is important to evaluate whether users understand the options that are available to them and the implications of their decision, given their (often) incomplete understanding of the technologies relevant to the privacy choice.

4.4.1 Objective Knowledge with Focused Attention

To better understand whether users can comprehend information provided in a privacy choice interface (either interruptive or on-demand), user study participants can be asked objective knowledge questions when it is presumed that their attention was focused on the privacy choice interface. This criterion can be assessed through user studies that involve privacy tasks, as well as self-report studies that ask participants to recall their experience with the privacy choice interface being evaluated. Koelle et al.'s study evaluating opt-in and opt-out gestures assessed objective knowledge by asking "What does the gesture shown in the video above mean to you" [31]. Evaluating for objective knowledge could also include asking if participants understand the privacy benefits and risks associated with different options, and if applicable, whether participants recognize whether a privacy choice is optional or mandatory.

4.4.2 Objective Knowledge with Unfocused Attention

It is also important to assess whether users understand the options available to them and implications of a decision made through interruptive privacy choice interfaces that they encounter when their attention is focused elsewhere in their interactions with a system. Similar to measuring awareness of a privacy choice described in 4.3.1, measuring objective knowledge with unfocused attention might require assigning participants to a distraction task, or having them recall their past experiences in a self-report study. Comparing objective knowledge when attention was focused on the privacy choice interface to when it was focused elsewhere may also help to reveal comprehension issues. For example, Pearman et al. asked participants about practices described in a HIPAA authorization they had encountered while trying to use a chatbot as part of a distraction task and later asked them to review the authorization again and revisit their answers [51].

4.4.3 Perceived Effort in Comprehending Choices

Similar to assessing the perceived effort to make a privacy choice (4.2.1), user study participants can be asked questions related to the perceived ease or difficulty in learning or comprehending the privacy choices. Similarly, this criterion can be assessed for both interruptive and on-demand interfaces after completing a study task that exposed them to the privacy choices. Alternatively, these questions can be asked about participants' prior experiences with a privacy choice interface outside of the study environment. Example prompts and measures to evaluate perceived learnability include: "what (if anything) was difficult to understand about the privacy choice interface" and items 5, 6, 7 and 10 on the SUS [35].

4.4.4 Estimated Effort in Comprehending Choices

Similar to the criteria described in 4.2.5 and 4.3.6, expert evaluation approaches can assess the difficulty in learning or comprehending a privacy choice interface. Such evaluations may assess whether particular types of users might have greater difficulty in learning or comprehending what the choice interface does, as well as what aid might be required to learn available choices. Furthermore, item 10 of the Nielsen heuristics also pertains to learnability [46].

4.5 User Sentiment

Different facets of user sentiment assess users' satisfaction with a privacy choice interface after they have had some exposure to it. This exposure may occur through a study task, or during their past interactions with a system. Evaluating for sentiment is applicable to both interruptive and on-demand privacy choice interfaces, and may be assessed through Likert measures accompanied with qualitative prompts.

4.5.1 Perceived Transparency & Control

This criterion assesses whether the privacy choice interface provides an appropriate level of transparency and control re-

lated to how user data is handled. Participants may be asked how transparent they feel the evaluated privacy choice interface is related to the use of their data, and to what extent they feel that it provides sufficient control over their data.

4.5.2 Subjective Knowledge

Assessing for subjective knowledge involves capturing users' interpretations of their ability to effectively use the privacy choice interface, as well as if they experience feelings of regret. Korff and Böhme used the TMC scale to measure participants' satisfaction, regret, and feelings of being overwhelmed after interacting with a privacy choice interface related to disclosure on a business networking website [32]. Example prompts and metrics related to subjective knowledge include to what extent participants feel informed about their choices, how capable they feel in making a decision, and how confident they are in their privacy choice (e.g., item 9 of SUS [35]).

4.5.3 Levels of Comfort and Trust

Ideally, privacy choice interfaces should empower users by providing control over their data. Thus it is important to evaluate whether after interacting with a privacy choice interface users are comfortable with how their data will be used, as well as to what extent they feel that their privacy decision will be honored. Mathur et al. argue that privacy choice interfaces should be evaluated on whether they are detrimental to the collective welfare [41]. In the context of privacy choice interfaces, dark patterns may result in a loss of trust or skepticism (e.g., in the company, in companies using similar privacy choice interfaces), and could contribute to feelings of resignation. Korff and Böhme also used the PCRT scale to measure participants' perceived comfort, risk, and trust in the privacy choice interface evaluated [32].

4.5.4 Investment in Decision-Making

This criterion pertains to whether the design of the privacy choice interface sufficiently motivates users to make an informed privacy decision. An example of prior work that assessed investment in decision-making is Cranor et al.'s user study that asked participants how likely they would to be to click on the do-not-sell icon and link texts being evaluated [12]. Other means of measuring investment include asking participants how carefully they considered their privacy choice and describing how they made their privacy decision.

4.6 Decision Reversal

For privacy choices to be usable, users need to be able to change their privacy choice decision, both immediately after an interaction with a privacy choice interface and, if applicable, at a later time through user settings offered through the website or app. This allows for users to correct an error they may have made in their initial privacy choice as well as circumstances in which users change their mind about how their data may be used or collected. The criteria for evaluating user ability & effort described in Section 4.2 related to making an initial privacy choice can be adapted to measure users' ability and effort in reversing their privacy decision (both immediately after making an initial decision and at a later point in time in which the choice interface or a settings page must be revisited). Similarly, those related to user awareness (Section 4.3 and user comprehension (Section 4.4) can be utilized to ensure that users can find and understand the information and processes that are part of reversing their privacy decision. Assessing for reversal through user studies involves assigning participants a privacy choice task in which they must undo or modify their initial privacy choice. This aspect of usability is applicable to both interruptive and on-demand interfaces.

4.7 Nudging Patterns

In contrast to the other usability aspects that are applicable to almost any type of user interface, evaluating for nudging patterns is especially relevant to contexts in which users are asked to give up something, such as their personal data. Privacy choice interfaces often exhibit dark patterns that nudge users to less privacy-protective outcomes to the benefit of the company. This usually occurs when privacy-protective options are made less salient or more cumbersome to use than the alternatives. Furthermore, legislation such as the General Data Protection Regulation (GDPR) and the California Privacy Rights Act (CPRA) make the use of dark patterns in privacy choice interfaces, particularly those related to consent, illegal [14, 48]. As such it is important for designers to be aware of the way they are nudging consumers and evaluate whether this nudging could be a dark pattern. In some contexts, it may even be appropriate for interfaces to nudge users to privacy-protective choices [1]. To evaluate interruptive and on-demand privacy choice interfaces for dark patterns, we propose criteria aligned to the normative perspectives described by Mathur et al. with regards to privacy [41].

4.7.1 Impact of Individual Welfare

Mathur et al. suggest measuring a "welfarist conception of privacy" [41]. In the privacy choice context, one such calculation is the financial value of the data disclosed because of a particular design pattern. User studies involving study tasks or self-reporting of data could also examine the proportion of users whose needs were not satisfied by a particular design. These measures could also highlight whether individual welfare could be improved with nudges toward privacy-protective choices. An example of prior work that has explored impact to individual welfare is Nouwen's et al.'s experiment that quantified the impact of different design elements in cookie consent interfaces on participants' consent decisions [47].

4.7.2 Unintended Societal Consequences

Another aspect of collective welfare is analyzing through expert evaluation approaches whether the privacy choice interface could lead to unintentional disclosure of personal information, and whether this could have negative societal-level impact. A prominent example is Facebook users unknowingly consenting to their data being shared with Cambridge Analytica, which used the data to influence global elections [42]. Gray et al.'s interaction criticism incorporated potential societal impact in a usability evaluation of cookie consent interfaces by including considerations such as "relevant business models and economic rationale, current and future role of technology, social acceptance or rejection of technology norms, agency of users and technology providers" [19].

4.7.3 Alignment with Regulatory Objectives

Expert evaluation approaches can also be used to ensure that designed privacy choice interfaces meet regulatory requirements. Both the GDPR and CPRA have provisions related to the usability of privacy choice interfaces, particularly to the consent of data collection [14, 48]. The CPRA explicitly bans dark patterns, defining them as "a user interface designed or manipulated with the substantial effect of subverting or impairing user autonomy, decision-making, or choice, as further defined by regulation." [48]. Prior empirical evaluations of consent notices have identified dark patterns that likely violate the spirit of GDPR and could potentially lead to regulatory penalties. Particularly Nowens et al. and Soe et al. provide a list of design criteria for cookie consent notices to evaluate for the presence of dark patterns and potential violations of the GDPR [47, 56]. This includes that consent be explicit (e.g., require a click from the user), consent must be as easy to withdraw or refuse as it is to give, and the privacy choice interface contain no pre-selected boxes for non-necessary purposes [47]. Other potentially violating design patterns are the absence of actual choices in the interface (e.g., instructions to change privacy choices are simply described in a notice text), choice toggles that are unlabelled, and not using antonyms of the consent option to label the option denying consent [56].

4.7.4 Individual Autonomy

Mathur et al. suggest evaluating to what degree an interface interferes with a user's ability to make "independent decisions" [41]. User study approaches can evaluate whether privacy choice interface designs lead users to choose certain privacy options over others by comparing privacy options selected through interfaces with suspected nudging patterns with those selected through other designs; cookie consent interface evaluations by Machuletz and Böhme [38] and Nouwens et al. [47] took this approach. Similarly, in some contexts it may be beneficial to evaluate whether interfaces utilizing reflective design better enable individual autonomy, as suggested by Terpstra et al [59]. Individual autonomy could also be evaluated through criteria that align with other evaluation objectives including: whether there is an option aligned with users' preferences available (4.1.1), whether users are able to choose their preferred option and the effort required (4.2.1, 4.2.2, 4.2.3), whether users are aware of the options available (4.3.1), whether users comprehend available options (4.4.1 and 4.4.2), and perceptions of autonomy (4.5.2 and 4.5.4).

5 Previous Privacy Choice Evaluations

This section presents an overview of a range of prior studies evaluating different types of privacy choice mechanisms. Though other work in this space may also be beneficial in informing the design of privacy choice interfaces, the studies described illustrate facets of the Privacy Choice Evaluation Framework through a variety of approaches. We focus our review on studies published over the past 10 years, with most published in the past five years.

A common privacy choice interface is related to allowing access to a specific hardware resource obtained from a device, like camera or location data. Previous studies have focused on user needs related to permission management in different contexts — including smartphone apps [26, 50], smart speakers [58], and smart glasses [13] — offering insights into the types of privacy controls that users desire. Other studies have uncovered limitations related to users' ability to use and comprehend existing permission management schemes [7, 55], or compared their usability to alternative approaches [60, 65]. Additionally, Bahirat et al. evaluated the impact of nudging on smart home privacy choices using data collected through hypothetical contextual scenarios, finding that defaults and framing of choices impact users' decision-making [4].

Interfaces that allow individuals to consent to different types of data processing are often used to meet legal requirements, such as those set by GDPR, Health Insurance Portability and Accountability Act (HIPAA), and Family Educational Rights and Privacy Act (FERPA). A growing body of work has explored the usability of cookie consent interfaces, finding that dark or nudging patterns that impact users' choices are prevalent in current interface designs [6, 19, 20, 47, 56, 61]. Others have explored the usability of consent interfaces used in other contexts. For example using an inspection-based approach, Khalil et al. found that students' ability to withdraw consent from Massive Open Online Course (MOOC) providers are limited due to lack of available options [30]. Additionally, Pearman et al. explored the usability of different health data disclosure authorization designs for a healthcare chatbot and argued for alternative approaches to capturing informed consent [51].

In contrast, other types of privacy choice interfaces allow users to opt out of the processing of their data, or to request deletion of their data. However, opt-out and deletion mecha-

nisms commonly used on websites and apps have been found to have usability issues related to awareness and ability & effort [18, 21, 24]. Other studies have explored visual icons as a potential means of increasing awareness of available opt-out choices through different user study designs, including participant inspection and assignment of a distraction task [12, 25]. Data deletion mechanisms have also been studied in the context of smart speakers; while users were found to be unaware of existing deletion options [40] the presence of available deletion mechanisms impacted users' trust in the system [9].

Privacy choice interfaces can also take the form of settings offered by a platform that users must typically seek out. Many studies have explored the usability of audience-related settings on content sharing and social media platforms, including user needs for audience control settings [16, 29, 57], tools that improve awareness of such settings [8], and users' ability to effectively use settings [28, 37, 39, 64]. Beyond audience settings, other studies have explored user needs for Facebook advertising controls through participant inspection [22], as well as the impact of social nudges on users' choice of Facebook privacy settings [33]. Outside of the social media context, Frik et al. collected self-reported data to explore the usability of smartphone privacy settings, highlighting issues of awareness and comprehension, among other usability issues [17].

Users can also make privacy choices through mechanisms decoupled from the original point of data collection. Past work utilizing participant inspection approaches has found that browser extensions may be effective in helping users become aware of available privacy opt-outs [5] and set their ideal privacy settings [34], but has highlighted that extensions themselves may be difficult for some users to configure [36]. Others have evaluated user needs for smartphone apps designed to aid privacy decision-making in different contexts [2, 11]. Furthermore, as traditional privacy choice mechanisms may be ill suited for some data collection scenarios by Internet of Things devices, others have explored the usability of alternative choice mechanisms [67] such as opt-out hand gestures [31, 68].

Altogether, this past work demonstrates the challenges of designing usable privacy choice mechanisms. Privacy decisions, such as content sharing, can be highly contextual [23]. Privacy choice interfaces must effectively communicate the scope of the privacy choice to allow users to make informed decisions [15]. The Privacy Choice Evaluation Framework presented can guide organizations in evaluating for aspects of usability pertinent to a particular privacy choice context.

6 Discussion

The Privacy Choice Evaluation Framework draws on evaluation approaches used in prior work to provide criteria to comprehensively evaluate the usability of privacy choice interfaces. The framework takes into account several considerations that make privacy choice interactions distinctive from interactions with other types of interfaces. The criteria provided in the framework can help guide organizations in evaluating new and existing privacy choice interface designs, which are necessary to support effective consumer privacy protection.

6.1 Additional Considerations

Privacy choice interactions differ from other interactions in that users are typically not trying to achieve a privacy goal when they interact with a system. Thus, the way they interact with privacy choice interfaces will be heavily impacted by their primary goal, such as to use a website or app. This is particularly relevant for interruptive privacy interfaces, such as cookie consent banners, which users may be inclined to quickly dismiss. This creates a tension between usability and privacy; while such interfaces may impede users in their primary goal and worsen the overall usability of a system, they can force users to make a privacy decision and offer an opportunity to select privacy-protective options that they would not have set otherwise. Furthermore, when evaluating privacy choice interfaces it is important to consider that users' behaviors and attitudes toward such interfaces are heavily influenced by their past experiences with similar privacy choices. Users may form expectations about where to find certain privacy choices and how they function [21]. Additionally, achieving meaningful privacy choice for some choice contexts in which users are overexposed to choice interfaces might require overcoming habituation and privacy fatigue [10].

The research methods described in the framework describe how general approaches to usability testing can be adapted to evaluate privacy choice interfaces. To ensure that meaningful privacy choice mechanisms are available to a broad population of internet users with differing abilities, evaluations utilizing these approaches should be performed in conjunction with accessibility assessments for which there are established frameworks [63]. In addition to users with disabilities, it is important to evaluate certain privacy choice interfaces with other vulnerable populations, such as marginalized racial groups or gender identities. Not only might these groups have specific privacy needs on a platform, the way they use existing privacy choice interfaces may differ from other users. An expert evaluation could provide an initial understanding of the usability of privacy choice interfaces for a special population. User studies with participants recruited from these special populations should be conducted to further this understanding.

6.2 Guidance for Organizations

A detailed example of how organizations can apply the Privacy Choice Evaluation Framework for their own usability evaluations is provided in the appendix. The same criteria could also be applied in studies that compare multiple privacy choice interface designs to identify which design elements are beneficial or detrimental to different usability aspects. In

selecting evaluation approaches, several factors related to the organization conducting the evaluation and the interface being evaluated should be considered. Here we describe a few such practical considerations.

Design Stage of the Privacy Choice Interface: An important factor that impacts which types of usability evaluations of a privacy choice interface are suitable is where in the design process the evaluation is being conducted. Ideally, evaluating the usability of a particular design would be integrated into an iterative design process with multiple research methods so that usability issues can be addressed prior to the interface being deployed. These usability assessments should build on each other. For example, a usability assessment in the ideation design phase may involve using qualitative methods, such as interviews or focus groups, to better understand users' needs in the context of the privacy choice interface. Expert evaluations, online surveys, experiments, and lab usability studies may be conducted with prototypes of the privacy choice interface to assess how well users' needs are met, as well as to what extent other usability aspects, including ability & effort, awareness, and comprehension, are achieved. Once a privacy choice interface is deployed, expert evaluations and field studies may be used to confirm that the usability of the final design is similar to results from previous usability testing.

Data Needed for Organizational Decisions: When considering the scope of possible research methods for assessments of privacy choice interfaces, it is necessary to prioritize which and what type of data are most important to capture from an organizational perspective. For example, some organizations may have additional requirements related to privacy choice that must be examined through a usability evaluation and thus focus more on a subset of the described usability aspects. Furthermore, organizations may differ in how they weigh and use different types of data in design decision-making. User studies that involve empirical data, such as field studies, online experiments, or lab usability studies, typically provide the best representation of how users may perceive or react to a particular design once it is deployed. However, user studies involving self-reported data may still provide enough of this insight to help organizations move forward with certain decisions. Expert evaluations can also aid in organizational decision-making, particularly in contexts where user feedback may not be helpful (e.g., new technologies where the average user may not be aware of all possible interaction paths).

Availability of Resources: Another important consideration in planning usability evaluations is the time, budget, and skill set of the evaluation team. While expert evaluations are typically less costly than user studies in terms of time and budget, they require evaluators with specific legal, design, or privacy expertise. User studies involving primarily quantitative data, such as surveys, can be deployed to a large number of participants (e.g., through online crowd-sourcing platforms) and analyzed in a short amount of time. Qualitative user studies may require more time for both data collection and analysis. Costs associated with user studies depend on factors such as the number of participants, length of the study, ease of recruiting qualified participants, amount of qualitative data to be analyzed, and depth of the analysis.

6.3 Limitations of Privacy Choice Usability

Better design of privacy choice interfaces, particularly those that allow users to decline data sharing just as easily as to agree to it, may be at odds with revenue-generating goals of a company. Though mounting consumer pressure should encourage companies to better privacy practices, it is still unclear whether this will translate to better consumer privacy protection. Privacy choice requirements in regulation, which include general requirements for usability, provide further incentive for companies to evaluate their privacy choice interfaces. While this framework could help organizations meet such usability requirements, and regulators to hold organizations accountable to better design practices, it is possible that interface designs that perform best in terms of usability (such as those that bundle certain privacy choices) would not be in full compliance with applicable legal requirements. Conversely, not all lawful designs of a privacy choice interface would perform well in meeting the framework's criteria.

Furthermore, even the most usable privacy choice interfaces place the burden of privacy management on users. In addition to privacy regulation, other mechanisms — such as technology supported decision-making and standardized privacy choice interfaces — are necessary to form a more effective consumer privacy protection framework. The Privacy Choice Evaluation Framework could serve as an initial step towards a more comprehensive implementation framework that could standardize interfaces for certain contexts. However until adoption of these privacy protection mechanisms becomes widespread, this framework provides immediately actionable guidance in improving privacy choice interfaces for users.

Acknowledgements

This research was supported in part by gifts from Facebook, the Carnegie Corporation of New York, and Innovators Network Foundation. We also would like to thank Alessandro Acquisti, Rebecca Balebako, Jessica Colnago, Yuanyuan Feng, Justin Hepler, Liz Keneski, Norman Sadeh, Hanna Schraffenberger, and Yixin Zou for their feedback on this work.

References

[1] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, et al. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys*, 50(3), 2017.

[2] Mamtaj Akter, Amy J. Godfrey, Jess Kropczynski, Heather R. Lipford, and Pamela J. Wisniewski. From parental control to joint family oversight: Can parents and teens manage mobile online safety and privacy as equals? *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW1), April 2022.

[3] Nick Babich. Tips to improve discoverability in UX, April 2020. https://xd.adobe.com/ideas/process/information-architecture/tips-to-improve-discoverability-in-ux/.

[4] Paritosh Bahirat, Martijn Willemsen, Yangyang He, Qizhang Sun, and Bart Knijnenburg. Overlooking context: How do defaults and framing reduce deliberation in smart home privacy decision-making? In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2021.

[5] Vinayshekhar Bannihatti Kumar, Roger Iyengar, Namita Nisal, Yuanyuan Feng, Hana Habib, Peter Story, Sushain Cherivirala, Margaret Hagan, Lorrie Cranor, Shomir Wilson, et al. Finding a choice in a haystack: Automatic extraction of opt-out statements from privacy policy text. In *Proceedings of The Web Conference*, 2020.

[6] Carlos Bermejo Fernández, Dimitris Chatzopoulos, Dimitrios Papadopoulos, and Pan Hui. This website uses nudging: Mturk workers' behaviour on cookie consent notices. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 2021.

[7] George Chalhoub, Martin J Kraemer, Norbert Nthala, and Ivan Flechais. "It did not give me an option to decline": A longitudinal analysis of the user experience of security and privacy in smart home products. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2021.

[8] Yi Chen, Mingming Zha, Nan Zhang, Dandan Xu, Qianqian Zhao, Xuan Feng, Kan Yuan, Fnu Suya, Yuan Tian, Kai Chen, et al. Demystifying hidden privacy settings in mobile apps. In *Proceedings of the IEEE Symposium on Security and Privacy (SP)*, pages 570–586. IEEE, 2019.

[9] Eugene Cho, S Shyam Sundar, Saeed Abdullah, and Nasim Motalebi. Will deleting history make Alexa more trustworthy? Effects of privacy and content customization on user experience of smart speakers. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2020.

[10] Hanbyul Choi, Jonghwa Park, and Yoonhyuk Jung. The role of privacy fatigue in online privacy behavior. *Computers in Human Behavior*, 81:42–51, 2018.

[11] Jessica Colnago, Yuanyuan Feng, Tharangini Palanivel, Sarah Pearman, Megan Ung, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. Informing the design of a personalized privacy assistant for the Internet of Things. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2020.

[12] Lorrie Faith Cranor, Hana Habib, Yaxing Yao, Yixin Zou, Alessandro Acquisti, Joel Reidenberg, Norman Sadeh, and Florian Schaub. CCPA opt-out icon testing–phase 2. Technical report, Office of the California Attorney General, 2020. https://www.oag.ca.gov/sites/all/files/agweb/pdfs/privacy/dns-icon-study-report-052822020.pdf.

[13] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno. In situ with bystanders of augmented reality glasses: Perspectives on recording and privacy-mediating technologies. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 2377–2386. ACM, 2014.

[14] European Parliament. Regulation (EU) 2016/679 of the European parliament and of the council, 2016. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679.

[15] Yuanyuan Feng, Yaxing Yao, and Norman Sadeh. A design space for privacy choices: Towards meaningful privacy control in the internet of things. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2021.

[16] Casey Fiesler, Michaelanne Dye, Jessica L Feuston, Chaya Hiruncharoenvate, Clayton J Hutto, Shannon Morrison, Parisa Khanipour Roshan, Umashanthi Pavalanathan, Amy S Bruckman, Munmun De Choudhury, et al. What (or who) is public? Privacy settings and social media content sharing. In *Proceedings of the Conference on Computer Supported Cooperative Work and Social Computing (CSCW)*, pages 567–580. ACM, 2017.

[17] Alisa Frik, Juliann Kim, Joshua Rafael Sanchez, and Joanne Ma. Users' expectations about and use of smartphone privacy and security settings. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2022.

[18] Stacia Garlach and Daniel Suthers. 'I'm supposed to see that?' AdChoices usability in the mobile environment. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS)*, 2018.

[19] Colin M Gray, Cristiana Santos, Nataliia Bielova, Michael Toth, and Damian Clifford. Dark patterns and the legal requirements of consent banners: An interaction criticism perspective. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2021.

[20] Hana Habib, Megan Li, Ellie Young, and Lorrie Faith Cranor. ''Okay, whatever": An evaluation of cookie consent interfaces. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2022.

[21] Hana Habib, Sarah Pearman, Jiamin Wang, Yixin Zou, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh, and Florian Schaub. "It's a scavenger hunt": Usability of websites' opt-out and data deletion choices. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2020.

[22] Hana Habib, Sarah Pearman, Ellie Young, Ishika Saxena, Robert Zhang, and Lorrie FaIth Cranor. Identifying user needs for advertising controls on facebook. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW1), April 2022.

[23] Hana Habib, Neil Shah, and Rajan Vaish. Impact of contextual factors on Snapchat public sharing. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2019.

[24] Hana Habib, Yixin Zou, Aditi Jannu, Neha Sridhar, Chelse Swoopes, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh, and Florian Schaub. An empirical analysis of data deletion and opt-out choices on 150 websites. In *Proceedings of the Symposium on Usable Privacy and Security (SOUPS)*. USENIX, 2019.

[25] Hana Habib, Yixin Zou, Yaxing Yao, Alessandro Acquisti, Lorrie Faith Cranor, Joel Reidenberg, Norman Sadeh, and Florian Schaub. Toggles, dollar signs, and triangles: How to (in)effectively convey privacy choices with icons and link texts. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2021.

[26] Qatrunnada Ismail, Tousif Ahmed, Apu Kapadia, and Michael K Reiter. Crowdsourced exploration of security configurations. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 467–476. ACM, 2015.

[27] ISO Technical Committee 159. Ergonomics of human-system interaction, March 2018. https://www.iso.org/standard/63500.html.

[28] Yousra Javed and Mohamed Shehab. Access control policy misconfiguration detection in online social networks. In *Proceedings of the International Conference on Social Computing (SocialCom)*, pages 544–549. IEEE, 2013.

[29] Dilara Kekulluoglu, Kami Vaniea, and Walid Magdy. Understanding privacy switching behaviour on Twitter. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2022.

[30] Mohammad Khalil, Paul Prinsloo, and Sharon Slade. The unbearable lightness of consent: Mapping MOOC providers' response to consent. In *Proceedings of the Conference on Learning at Scale (L@S)*. ACM, 2018.

[31] Marion Koelle, Swamy Ananthanarayan, Simon Czupalla, Wilko Heuten, and Susanne Boll. Your smart glasses' camera bothers me! Exploring opt-in and opt-out gestures for privacy mediation. In *Proceedings of the Nordic Conference on Human-Computer Interaction (NordiCHI)*, pages 473–481, 2018.

[32] Stefan Korff and Rainer Böhme. Too much choice: End-User privacy decisions in the context of choice proliferation. In *Proceedings of the Symposium On Usable Privacy and Security (SOUPS)*, pages 69–87, 2014.

[33] Isadora Krsek, Kimi Wenzel, Sauvik Das, Jason I Hong, and Laura Dabbish. To self-persuade or be persuaded: Examining interventions for users' privacy setting selection. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2022.

[34] Oksana Kulyk, Peter Mayer, Melanie Volkamer, and Oliver Käfer. A concept and evaluation of usable and fine-grained privacy-friendly cookie settings interface. In *Proceedings of the International Conference on Trust, Security And Privacy in Computing and Communications/International Conference on Big Data Science and Engineering (TrustCom/BigDataSE)*, pages 1058–1063. IEEE, 2018.

[35] Page Laubheimer. Beyond the NPS: Measuring perceived usability with the SUS, NASA-TLX, and the single ease question after tasks and usability tests, February 2018. https://www.nngroup.com/articles/measuring-perceived-usability/.

[36] Pedro Giovanni Leon, Blase Ur, Richard Shay, Yang Wang, Rebecca Balebako, and Lorrie Faith Cranor. Why Johnny can't opt out: A usability evaluation of tools to limit online behavioral advertising. In *Proceedings of*

the Conference on Human Factors in Computing Systems (CHI). ACM, 2012.

[37] Yabing Liu, Krishna P Gummadi, Balachander Krishnamurthy, and Alan Mislove. Analyzing Facebook privacy settings: User expectations vs. reality. In *Proceedings of the Internet Measurement Conference (IMC)*, pages 61–70. ACM, 2011.

[38] Dominique Machuletz and Rainer Böhme. Multiple purposes, multiple problems: A user study of consent dialogs after GDPR. *Proceedings on Privacy Enhancing Technologies*, 2020(2):481–498, 2020.

[39] Michelle Madejski, Maritza Johnson, and Steven M Bellovin. A study of privacy settings errors in an online social network. In *Proceedings of the International Workshop on Security and Social Networking (SESOC)*, pages 340–345. IEEE, 2012.

[40] Nathan Malkin, Joe Deatrick, Allen Tong, Primal Wijesekera, Serge Egelman, and David Wagner. Privacy attitudes of smart speaker users. *Proceedings on Privacy Enhancing Technologies*, 2019(4):250–271, 2019.

[41] Arunesh Mathur, Jonathan Mayer, and Mihir Kshirsagar. What makes a dark pattern...dark? Design attributes, normative considerations, and measurement methods. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2021.

[42] Sam Meredith. Here's everything you need to know about the Cambridge Analytica scandal. *CNBC*, March 2018. https://www.cnbc.com/2018/03/21/facebook-cambridge-analytica-scandal-everything-you-need-to-know.html.

[43] Justin Mifsud. Usability metrics – a guide to quantify the usability of any system. https://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/.

[44] Peter Morville. User experience design, June 2004. http://semanticstudios.com/user_experience_design/.

[45] Jakob Nielsen. Usability 101: Introduction to usability, January 2012. https://www.nngroup.com/articles/usability-101-introduction-to-usability/.

[46] Jakob Nielsen. 10 usability heuristics for user interface design, November 2020. https://www.nngroup.com/articles/ten-usability-heuristics/.

[47] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2020.

[48] Office of the California Attorney General. The California Privacy Rights and Enforcement Act of 2020, 2019. https://oag.ca.gov/system/files/initiatives/pdfs/19-0017%20%28Consumer%20Privacy%20%29.pdf.

[49] OneTrust. DHL increases CMP opt-in rates with A/B testing and OneTrust PreferenceChoice, June 2021. https://www.cookiepro.com/wp-content/uploads/2021/06/20210421-OneTrust-DHL-CS-US-Digital.pdf.

[50] Sameer Patil, Roman Schlegel, Apu Kapadia, and Adam J Lee. Reflection or action? How feedback and control affect location sharing decisions. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 101–110. ACM, 2014.

[51] Sarah Pearman, Eleanor Young, and Lorrie Faith Cranor. User-friendly yet rarely read: A case study on the redesign of an online HIPAA authorization. *Proceedings on Privacy Enhancing Technologies*, 2022(3), 2022.

[52] Whitney Quesenbery. Balancing the 5Es: Usability. *Cutter IT Journal*, February 2004. http://whitneyquesenbery.com/articles/5es-citj0204.pdf.

[53] John A Rothchild. Against notice and choice: The manifest failure of the proceduralist paradigm to protect privacy online (or anywhere else). *Cleveland State Law Review*, 66, 2017.

[54] Florian Schaub and Lorrie Faith Cranor. Usable and useful privacy interfaces. In Travis Breaux, editor, *An Introduction to Privacy for Technology Professionals*, pages 176–299. IAPP, 2020.

[55] Muhammad Umair Shah, Umair Rehman, Farkhund Iqbal, Fazli Wahid, Mohammed Hussain, and Ali Arsalan. Access permissions for Apple Watch applications: A study on users' perceptions. In *Proceedings of the International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*. IEEE, 2020.

[56] Than Htut Soe, Oda Elise Nordberg, Frode Guribye, and Marija Slavkovik. Circumvention by design - dark patterns in cookie consent for online news outlets. In *Proceedings of the Nordic Conference on Human-Computer Interaction (NordiCHI)*, 2020.

[57] Katherine Strater and Heather Richter. Examining privacy and disclosure in a social networking community. In *Proceedings of the Symposium on Usable Privacy and Security (SOUPS)*, 2007.

[58] Madiha Tabassum, Tomasz Kosiński, Alisa Frik, Nathan Malkin, Primal Wijesekera, Serge Egelman, and Heather Richter Lipford. Investigating users' preferences and expectations for always-listening voice assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, 3(4), 2019.

[59] Arnout Terpstra, Paul Graßl, and Hanna Schraffenberger. Think before you click: How reflective patterns contribute to privacy. In *Proceedings of the What Can CHI Do About Dark Patterns Workshop*. ACM, 2021.

[60] Lynn Tsai, Primal Wijesekera, Joel Reardon, Irwin Reyes, Serge Egelman, David Wagner, Nathan Good, and Jung-Wei Chen. Turtle Guard: Helping android users apply contextual privacy preferences. In *Proceedings of the Symposium on Usable Privacy and Security (SOUPS)*, 2017.

[61] Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. (Un)informed consent: Studying GDPR consent notices in the field. In *Proceedings of the Conference on Computer and Communications Security (CCS)*, pages 973–990. ACM, 2019.

[62] Kami Vaniea, Lujo Bauer, Lorrie Faith Cranor, and Michael K Reiter. Studying access-control usability in the lab: Lessons learned from four studies. In *Proceedings of the Workshop on Learning from Authoritative Security Experiment Results*, pages 31–40, 2012.

[63] W3C Web Accessibility Initiative. Web content accessibility guidelines (WCAG) 2.1, 2018. https://www.w3.org/TR/WCAG21/.

[64] Yang Wang, Liang Gou, Anbang Xu, Michelle X Zhou, Huahai Yang, and Hernan Badenes. Veilme: An interactive visualization tool for privacy configuration of using personality traits. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 817–826, 2015.

[65] Primal Wijesekera, Joel Reardon, Irwin Reyes, Lynn Tsai, Jung-Wei Chen, Nathan Good, David Wagner, Konstantin Beznosov, and Serge Egelman. Contextualizing privacy decisions for better prediction (and protection). In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2018.

[66] Chauncey Wilson. *User Interface Inspection Methods: A User-Centered Design Method*. Newnes, 2013.

[67] Yaxing Yao, Justin Reed Basdeo, Smirity Kaushik, and Yang Wang. Defending my castle: A co-design study of privacy mechanisms for smart homes. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. ACM, 2019.

[68] Yaxing Yao, Huichuan Xia, Yun Huang, and Yang Wang. Privacy mechanisms for drones: Perceptions of drone controllers and bystanders. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pages 6777–6788. ACM, 2017.

A Using the Evaluation Framework

Table 1 provides a summary of the criteria included in the Privacy Choice Evaluation Framework. The mapping of the criteria to usability aspects, evaluation approaches, and interface timings was informed by prior work and standard HCI practices. Organizations and other researchers can use this table as a reference when designing evaluation studies. First, researchers should identify their study goals, or the usability aspect(s) that they want to explore in their usability study. Then researchers should identify an evaluation approach that is suited for addressing their study goals, taking into account the considerations described in Section 6.2. For more comprehensive usability evaluations, researchers may choose to incorporate multiple evaluation approaches into their study. Last, the researchers should select the criteria that are appropriate for the particular privacy choice context, taking into account the timing of the choice interface being evaluated.

In prior work, we used the Privacy Choice Evaluation Framework to assess the usability of 12 cookie consent banner variants, touching on a large fraction of the criteria [20]. This may serve as a useful example for understanding how the framework might be used.

Here we describe a scenario where you might want to do a fairly comprehensive evaluation. Imagine you are working for a company developing a new skincare app that allows users to take photographs of their skin, get recommendations for skincare products, get referrals to dermatologists and skincare professionals, and discuss skincare issues with other users.

As you begin developing the app, you conduct focus groups to understand the interests of potential users. During this phase, it would be a good idea to also focus on users' privacy **needs** by conducting interviews, focus groups, or surveys to uncover *users' privacy objectives*, including the types of privacy choices they would like to have and whether there are any special requirements for this population of users — who might include acne-prone teenagers under age 18 and people who suffer from chronic skin conditions or are experiencing skin problems as a side effect of treatments for other conditions. Here it would be useful to find out whether course-grained controls over data sharing would meet users' needs or if (some) users would appreciate finer-grained controls over the type of data to be shared, with whom it is shared, or other privacy objectives. In this phase you may discover that some users have little sensitivity about discussing certain types of skincare concerns and are interested in getting advertisements and discounts on relevant products, while other users are interested in getting advice from experts and other users with

their condition, but are concerned about being identified as a person with a particular condition and do not want to receive related advertisements.

As low- to mid-fidelity prototypes of the app are developed (such as static wire frames or interactive prototypes), user studies can probe other aspects of the framework with participants representative of those expected to use the app, including the special populations identified. For example, a lab or online study might present prototypes to participants, ask them to step through some typical non-privacy tasks, and then ask them about what information they believe is being shared and what privacy choices are available to probe **awareness**, particularly *awareness of choice existence*. Then participants might be directed to *find the privacy choice* interface and *make privacy choices* they would like to have to evaluate their **ability & effort** using an **on demand** privacy choice interface. Researchers may want to ask participants about their *privacy objectives* and *privacy intentions* to confirm that choices meet the participant's needs and to see whether the choices the participant made align with their stated intentions. Participants might also be asked questions related to **comprehension** to assess their *objective knowledge* of the privacy choices available and what they do. Finally, participants may be asked questions pertaining to **sentiment**, for example to assess their *investment in decision-making*, *perceived levels of transparency and control over how their data will be used*, *self-efficacy in using the choice interface*, and *comfort and trust in the company's handling of their data*.

If the app includes a feature with an **interruptive** privacy choice interface, such as a prompt for the user to immediately make a decision about whether an uploaded photograph will be shared, users should be asked to perform a task that triggers the interruption and then similar evaluations should be conducted as with the on-demand interface. Here participants might be asked **comprehension** questions to assess their *objective knowledge* of the privacy choices available and what they do, both after completing the choice task with the choices no longer visible on screen (*unfocused attention*), and when revisiting the choice interface (*focused attention*). The evaluation of **user sentiment**, such as *investment in decision-making*, here is even more relevant than in the on-demand task, as it allows an assessment of whether participants were trying to make a meaningful decision at the time the choice appeared or just swatting the prompt away. To evaluate **decision reversal**, participants may be asked what they would do if they wanted to change their privacy decision. This user study data might also be helpful for evaluating for potential **nudging patterns**, particularly whether the interface designs hinders *individual welfare* or *individual autonomy* .

A usable privacy **expert** may evaluate the privacy choice interface for **user needs, ability & effort, awareness, and comprehension**. An expert may examine *interface completeness* and *interface accuracy* related to the needs uncovered in prior evaluations, and also *estimate the effort needed to make a privacy choice, users' abilities to find the privacy choice, and comprehension of the choices*. A privacy legal expert might evaluate for potential **nudging patterns** by examining *alignment with regulatory objectives*, including any relevant laws concerning sensitive health information or children's privacy, as well as any *unintended societal consequences* of the interface.

As app development proceeds, some of these studies would be repeated with higher fidelity prototypes and eventually the finished app. Where potential problems are uncovered, alternate interfaces might be tested and compared. In some cases a very narrow study might be done to focus on a specific problem, for example, if users are having trouble understanding a particular privacy choice, an online survey might just probe comprehension of alternative ways of describing that choice. Once improved language is identified it should then be tested in the full app context.

The number of study participants and number of rounds of iteration will vary depending on the complexity of the app, number of problems surfaced in the initial studies, resources available, and objectives of the app developers. Different levels of rigor are needed for published academic papers than for internal testing. However, a company that is under regulatory scrutiny, trying to hold itself up as a privacy role model, or planning to publish the results of its internal testing may engage in more rigorous testing than a company that just wants to do enough testing to avoid major privacy pitfalls.