The Recipe for the Perfect Batch: Assessing New Cookie Category Terms

Adhishree Abhyankar
Carnegie Mellon University

Lydia Barit
Carnegie Mellon University

Soha Jiwani
Carnegie Mellon University

Rachna Sasheendran
Carnegie Mellon University

Mireya Sotelo
Carnegie Mellon University

Lorrie Cranor
Carnegie Mellon University

Abstract

In 2012, the UK International Chamber of Commerce (ICC) released a Cookie Guide containing standardized cookie terms that are now used all around the web [2]. However, even though users frequently see these terms, recent research shows that they may not actually understand them [5]. Our multi-stage study surfaces alternate terminology and assesses the comprehension of these new terms against existing terms. We gathered a set of alternate cookie category terms through a pilot focus group and extracted terms from two popular Consent Management Platforms (CMPs). We then conducted a 100-participant survey to filter the terms that best described the initial cookie categories. Finally, we conducted a 120-participant survey to measure the comprehensibility of these filtered terms. We found that in three of the four cookie categories, alternate terms provided better comprehensibility than one or more of their original counterparts. These findings suggest that a review of this guide should be undertaken to ensure that users are presented with understandable cookie terminology when interacting with cookie consent interfaces.

1 Introduction

In 2012, the United Kingdom's International Chamber of Commerce released the ICC UK Cookie guide, a guide designed to "help website operators to provide information to users in language they can understand" through the use of standardized web cookie terminology [2]. This guide divides cookies into four categories: *Strictly Necessary Cookies*, *Performance Cookies*, *Functionality Cookies*, and *Targeting/Advertising Cookies*. These terms have been adopted by popular CMPs, including OneTrust [3].

While prior research has focused more on the visual aspects of cookie consent interfaces from legal, regulatory, usability, and deceptive practices angles, little work has been done to assess the understandability of these decade-old terms. Recently, research has been published that shows that most people don't understand current cookie category terminology [5].

Our project aims to fill this gap by conducting a study regarding the comprehensibility of the current cookie terminology, as well as coming up with alternate terms that may be easier to understand. Specifically, our two research questions are as follows:

1. What terms or phrases could better describe the different types of cookie categories?

2. Do these terms or phrases lead to a better user understanding of cookie categories than the existing set of terms?

2 Related Work

Prior research has shown that users do not fully understand what different types of cookies do. Habib et al., through the results of a 1,109 participant definition-selection multiple choice survey, showed that only 47.6% of participants picked the correct definition for performance cookies, and only 16% of participants picked the correct definition for functional cookies [5]. Not only do users not understand what different types of cookies do based on the terms that describe them, but they also lack an understanding of the purposes of such cookies [7]. Miyazaki's findings are also consistent with those from Pinto et al. that show that many users accept cookies without fully understanding their purposes [8]. This lack of knowledge leaves users, especially novice internet users, overwhelmed when confronted with cookie consent interfaces [7,8]. Utz et al. explain that the amount of information given to the consumer through these interfaces dictates the insight users get about cookies that are used [9]. Often, websites provide users with too many cookie categories to opt-in or opt-out of without much information, not allowing users to fully understand to what degree they have privacy on that website [9].

Through our research, we found a lack of clear, concise naming conventions for categories that allows users to know exactly what they are consenting to when accepting cookies.

Thus, we build on this prior research by finding more intuitive and understandable terms to describe different categories of cookies, helping users to make more informed decisions when confronted with cookie consent interfaces.

3 Methodology

The following section describes the construction and implementation of our two surveys. All of our study's components were reviewed and approved by Carnegie Mellon University's Institutional Review Board (IRB).

3.1 Survey 1

We conducted our first survey with the goal of finding cookie category terms that were apt descriptors of the definitions specified in the ICC UK Cookie Guide. For each definition, we presented the user with a currently-used cookie category term along with new terms that were formulated during a pilot focus group and gathered from two popular CMPs (Cookiebot and OneTrust). Participants were asked how well they believed each of the terms matched the definition, and why they preferred certain terms over others for a specific definition.

We recruited 100 participants via Prolific for our first user survey, hosted on Qualtrics. Participants had to be 18 years or older, fluent in English, and situated in the US. The users were asked to provide their informed consent and asked a series of demographic questions. The survey was estimated to take an average of 5 minutes to complete. Participants were compensated a total of $1.00 for completing the survey.

3.1.1 Survey Questions

Our survey included cookie category term rating questions with Likert scale responses for the purpose of examining the terms that best match the definitions for each cookie category. The list of terms to be analyzed under each category definition included terms from our pilot focus group as well as the existing cookie category terms from Cookiebot [1] and OneTrust [3]. Since the cookie category definitions are written in British English and all of our participants were based in the US, we Americanized the definitions by replacing certain letters and words. We were cautious to avoid including suggested cookie category terms that may be too technical for the average user to understand while picking terms to include in the first survey. Each term rating question was followed by two open ended questions. The first question collected the reasons behind the participant's opinion as to why they thought some terms seemed to match or not match the respective cookie category definition. The second free-response question asked the participants to suggest terms that they believed would better describe the cookie category according to the definition given. The complete survey provided to the participants can be found in Appendix Section 7.6.

3.1.2 Coding Free-responses

One of our free-response questions asked participants for their reasoning as to why they thought some terms seemed to match or not match the definition. We used the method of emergent coding and initially analyzed 60 responses as a group, pulled out the common patterns we observed in these responses, and created our code-book based on these patterns. Once a comprehensive code-book was developed (see Appendix Figure 1), three members of our group individually coded each of the responses. After the three coders finished individually coding the responses, the three coders then cross-checked the assigned codes. When there were conflicting assigned codes, the coders discussed and settled on a final code for the response.

3.2 Survey 2

The second survey focused on measuring the comprehension of the terms that emerged from Survey 1 as well as original terms. We then compared the comprehension levels of all terms in a given cookie category to the original term to see if any of them were significantly better comprehended by users.

120 survey participants aged 18 and over, fluent in English, and based in the United States were recruited via Prolific. The survey was made to take 5 minutes to complete, and participants were given a total compensation of $1.00. Additionally, we excluded participants from the Survey 1 from taking Survey 2 to avoid primed participants.

3.2.1 Survey Questions

The survey questions were structured such that participants had to choose the correct definition from a list of definitions provided for each term we wanted to test. The short definitions used for each multiple choice option were taken from Habib et al. [5]. For terms belonging to the same category, the same definition options were used with only one correct option, (i.e., we test comprehension levels through correctness). To avoid learning effects as much as possible, we randomized the order in which participants saw questions. A copy of the survey questions can be found in Section 7.7 of the Appendix.

4 Results

4.1 Demographics

Survey 1 Of the 100 responses that we received, 10 were excluded from our analysis. We based our exclusion criteria on the completeness of open-ended questions, whether open-ended responses matched what was selected in immediately preceding Likert scale questions, and whether or not a participant selected the same Likert scale rating across all questions in the survey. Our survey's population skewed heavily female and young, possibly due an influx of young female registrants

on Prolific after a TikTok video went viral [4, 5]. Participants most commonly rated their familiarity with privacy concepts as a 3 out of 5 (44.4%) followed by a 4 out of 5 (32.2%).

Survey 2 Our survey population of 120 skewed young, as was the case with the population in Survey 1. The majority of participants had a Bachelor's degree (35.8%) or some college education but no degree (30.0%). Additionally, participants most commonly rated their familiarity with privacy concepts as a 3 out of 5 (45.8%) followed by a 4 out of 5 (27.5%). We did not exclude any responses.

4.2 Survey 1 Free-response Insights

We noticed that participants primarily provided reasons for preferring some terms over others or reasons for not preferring a particular term, shaping our first level of coding to differentiate between "positive" and "negative" responses. Participants reacted positively to terms that seemed to be intuitive, comprehensible, and concise, while reacting negatively to those appearing to be confusing or wordy. Confusion mostly stemmed from vague terms or terms that seemed to be in conflict with their definition. Our second free-response question asked participants to provide suggestions for the terms that could better match these cookie category definitions. While we did not test the comprehension of these terms, we have provided a list in Appendix Table 4. In the future, the comprehension of these terms could be tested to expand on this study.

4.3 Survey 1: Term Selection

In order to surface the top alternative terms in each of the four cookie categories, we added up the frequencies of the Likert scale responses for each term from Survey 1. Following this, we calculated a "Top 2" sum and a "Bottom 2" sum for each term. Calculating "Top 2" sums involved adding the frequency of "Extremely well" and "Very well" responses for each term, and calculating "Bottom 2" sums involved adding the frequency of "Not well at all" and "Slightly well" responses for each term. We chose to use this measurement in order extract terms that were both well liked and not widely disliked (i.e., finding terms with relatively high Top 2 sums and relatively low Bottom 2 sums). As a result, the following terms were selected to be tested for comprehension in Survey 2:

- **Strictly Necessary Cookies**: Functional Cookies, Necessary Cookies, Website Function Cookies

- **Functionality Cookies**: Personalized Experience Cookies, Preferences Cookies

- **Performance Cookies**: Analytics Cookies

- **Advertising/Targeting Cookies**: Marketing Cookies, Third Party Advertising Cookies, Personalized Advertising Cookies

The bold text above represents the original term followed by alternate terms to test. A summary of all response frequencies and Top 2 and Bottom 2 sums for all terms tested can be found in Appendix Table 1.

4.4 Term Comparison

To measure term comprehension from Survey 2 responses, we calculated the proportion of correct answers for each comprehension question. Within each cookie category, we conducted paired t tests to compare the comprehension of all alternative terms to original terms under the following hypotheses:

$$H_0: \ \mu_O - \mu_A = 0$$
$$H_1: \ \mu_O - \mu_A > 0,$$

where μ_O indicates the proportion of correct answers among participants when they were presented with the original term and μ_A indicates the same for each alternate term listed in Section 4.3.

Since we conducted multiple hypothesis tests, we corrected for Type I errors using the Holm correction [6]. For all of our tests, we employ an α level of 0.05.

We also examined the incorrect answer choices for each term in each category to see whether a large proportion of participants incorrectly understood a term to mean something else.

4.4.1 Category 1: Strictly Necessary Cookies

We did not notice a significant difference in comprehensibility between Strictly Necessary($\mu = 0.733$) and the alternative terms tested for this category (Functional ($\mu = 0.708$, $p = 0.991$); Necessary ($\mu = 0.675$, $p = 0.612$); Website Function ($\mu = 0.817$, $p = 1.000$)). However, since coding results from our first survey confirmed that users prefer concise terms, future work should not ignore testing these alternatives. We did not notice any clustering around any particular incorrect answers.

4.4.2 Category 2: Performance Cookies

We saw a significant difference in comprehension when users were presented with Analytics Cookies ($\mu = 0.858$, $p = 8.9e - 14$) as the cookie term relative to Performance Cookies Performance($\mu = 0.425$). We did not notice any clustering around any particular incorrect answers.

4.4.3 Category 3: Functionality Cookies

We saw a significant difference in comprehension when users were presented with either of the alternative terms, Personalized Experience ($\mu = 0.892$, $p < 2e - 16$) and Preferences ($\mu = 0.800$, $p < 2e - 16$), relative to Functionality($\mu = 0.167$). In analyzing participants' incorrect answer choices, we also

found that 78 out of 120 participants (65%) incorrectly stated that Functionality Cookies are "cookies that are needed for the website to work properly." For other terms in this category, we did not see clusters of this magnitude forming at any particular incorrect answer.

4.4.4 Category 4: Targeting/Advertising Cookies

Since this category consists of two original terms from the ICC UK Cookie Guide, we present results of hypothesis tests that consider Advertising Cookies as the original term and Targeting Cookies as the original term. In this way, we can discern whether one is comparatively more comprehended that the other.

Original term: Advertising Cookies We did not notice a significant difference in comprehensibility between Advertising ($\mu = 0.967$) and the alternative terms tested for this category (Marketing ($\mu = 0.892$, $p = 1.000$); Third Party Advertising ($\mu = 0.833$, $p = 1.000$); Personalized Advertising ($\mu = 0.983$, $p = 1.000$)). We discuss the test between Advertising Cookies and Targeting Cookies below.

Original term: Targeting Cookies We saw a significant difference in comprehension when users are presented with any of the new terms (Marketing ($\mu = 0.892$, $p = 3.7e-08$), Third Party Advertising ($\mu = 0.833$, $p = 4.1e-05$), Personalized Advertising ($\mu = 0.983$, $p = 2.5e-14$)) and the co-original term (Advertising ($\mu = 0.967$, $p = 4.0e-13$)) relative to Targeting ($\mu = 0.650$).

In analyzing participants' incorrect answer choices, we also found that 29 out of 120 participants (24.2%) incorrectly stated that Targeting Cookies are "cookies that are needed for determining your location." For other terms, we did not see clusters of this magnitude forming at any particular incorrect answer.

5 Discussion

From our results, we saw that alternate terms for Categories 2 and 3 were significantly more comprehensible than the existing terms. This is consistent with findings from prior research done by Habib et al., where it was found that these two cookie category terms in particular were the most misunderstood among the study's participants [5]. Further, the majority of participants in our study thought that Functionality Cookies were "cookies that are needed for the website to work properly." Our results present conflicting insights for Category 4 (Advertising/Targeting Cookies) cookie terms. While alternate terms did not significantly increase comprehension when "Advertising Cookies" was tested as the original term, the opposite occurred when "Targeting Cookies" was tested as the original term. Future work should evaluate whether removing "Targeting Cookies" from the guide would better standardize cookie terminology and improve user understanding. Future work should also continue to evaluate all alternate

terms for this category as discussed in Section 6. Lastly, our study found that alternate terms for Category 1 (Strictly Necessary Cookies) were not significantly more comprehensible compared to the original term. However, based on our Survey 1 coding results, future work should focus on whether a more concise term would be a better alternative.

6 Conclusion and Future Work

Our multi-staged study surfaced alternate cookie terms and assessed their comprehensibility alongside a set of terms that are currently in use. Generally, we found that users prefer terms that are intuitive, comprehensible, and concise. Additionally, in three of the four cookie categories, alternate terms provided better comprehensibility than their original counterparts. These findings suggest that a review of the ICC UK Cookie Guide should be undertaken to ensure that users are presented with understandable cookie terminology when interacting with cookie consent interfaces.

While our study provides foundational insight into the weaknesses of existing cookie terminology, future scholarship is required to understand which term(s) in each category are best, including testing those that were suggested in Survey 1. In tandem with this work, focus should be placed on revising the ICC UK Cookie Guide's cookie category definitions, the definitions that will ultimately be described by the aforementioned terms since many participants, without provocation, pointed out how confusing they were. Lastly, we assessed the comprehension of original and alternate cookie terms in a survey format rather than *in situ*. Conducting a follow up study that assesses these terms as they are interacted with in a realistic cookie interface may provide more nuanced insights around comprehension.

References

[1] Cookie texts and cookie messages in the age of the privacy paradox. Cookiebot CMP. https://www.cookiebot.com/en/cookie-texts/.

[2] ICC UK Cookie guide. Technical report, International Chamber of Commerce, London, UK, November 2012.

[3] Cookie Notice. OneTrust. https://www.onetrust.com/cookie-policy, 2020.

[4] Charalambides, Nick. We recently went viral on tiktok - here's what we learned, 2021.

[5] Hana Habib, Megan Li, Ellie Young, and Lorrie Faith Cranor. "okay, whatever": An evaluation of cookie consent interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.

[6] Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

[7] Anthony D. Miyazaki. Online privacy and the disclosure of cookie use: Effects on consumer trust and anticipated patronage. *Journal of Public Policy & Marketing*, 27(1):19–33, 2008.

[8] Pedro Pinto, Romeu Lages, and Manuel Au-Yong Oliveira. *Web Cookies: Is There a Trade-off Between Website Efficiency and User Privacy?*, pages 713–722. 06 2020.

[9] 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 2019 ACM SIGSAC Conference on Computer and Communications Security*, CCS '19, page 973–990, New York, NY, USA, 2019. Association for Computing Machinery.

7 Appendix

7.1 Survey 1 Cookie Term Frequencies and Selection Criteria

Original Term	Terms Assessed	NW	SW	MW	VW	EW	Top 2	Bottom 2	Chosen?
Strictly Necessary Cookies	Functional Cookies	3	10	17	38	22	60	13	Yes
	Necessary Cookies	2	9	20	32	27	59	11	Yes
	Status Cookies	46	23	17	4	0	4	69	No
	Strictly Necessary Cookies	11	13	20	26	20	46	24	Original
	User State Cookies	32	35	16	7	0	7	67	No
	Website Function Cookies	1	4	22	41	22	63	5	Yes
Performance Cookies	Analytics Cookies	2	6	19	36	27	63	8	Yes
	Interaction Statistic Cookies	10	28	29	20	3	23	38	No
	Metric Cookies	13	23	20	29	5	34	36	No
	Performance Cookies	14	21	24	24	7	31	35	Original
	Site Statistic and Error Handling Cookies	11	14	23	29	13	42	25	No
	Statistics Cookies	4	20	25	34	7	41	24	No
	Visitation and Error Checking Cookies	5	22	24	29	10	39	27	No
Functionality Cookies	Consistent Experience Cookies	13	16	27	24	10	34	29	No
	Ease-of-access Cookies	10	22	20	26	12	38	32	No
	Ease-of-navigation Cookies	10	18	30	24	8	32	28	No
	Functional Cookies	14	29	27	18	2	20	43	No
	Functionality Cookies	20	24	32	14	0	14	44	Original
	Personalized Experience Cookies	6	6	13	39	26	65	12	Yes
	Preferences Cookies	6	4	16	34	30	64	10	Yes
	Session Enhancing Cookies	18	19	32	17	4	21	37	No
Advertising/ Targeting Cookies	Advertising Cookies	1	2	6	34	47	81	3	Original
	Individual Behavioral Advertising Cookies	10	18	24	27	11	38	28	No
	Marketing Cookies	2	4	16	39	29	68	6	Yes
	Targeting Cookies	6	16	16	32	20	52	22	Original
	Third Party Advertising Cookies	5	6	18	35	26	61	11	Yes
	Personalized Advertising Cookies	4	15	14	35	22	57	19	Yes

Table 1: Frequency counts of Likert scale ratings for each term. NW = Not well at all; SW = Slightly well, MW = Moderately well, VW = Very well, EW = Extremely well."Chosen?" indicates whether the term met the frequency requirements to be chosen for assessment in Survey 2. All original terms were tested in Survey 2.

7.2 Survey Demographics

7.2.1 Survey 1

Age		Gender		Education		Privacy Concepts Familiarity	
18 - 24	18.9%	Male	17.8%	High school diploma or GED	3.3%	1	3.3%
25 - 34	35.6%	Female	77.8%	Some college education but no degree	16.7%	2	12.2%
35 - 44	20.0%	Non-Binary	3.3%	Associate's Degree	15.6%	3	44.4%
45 - 54	11.1%	Not specified	1.1%	Bachelor's Degree	43.3%	4	32.2%
55 - 64	6.7%			Master's Degree	17.8%	5	7.8%
65 - 74	6.7%			Doctorate	2.2%		
75 - 84	1.1%			Self-specified	1.1%		

Table 2: Participants demographics in Survey 1.

7.2.2 Survey 2

Age		Gender		Education		Privacy Concepts Familiarity	
18 - 24	19.2%	Male	50.0%	Less than high school diploma	1.7%	1	4.2%
25 - 34	35.8%	Female	46.7%	High school diploma or GED	7.5%	2	14.2%
35 - 44	25%	Non-Binary	2.5%	Some college education but no degree	30.0%	3	45.8%
45 - 54	10.8%	Self described	0.8%	Associate's Degree	8.3%	4	27.5%
55 - 64	8.3%			Bachelor's Degree	35.8%	5	8.3%
65 - 74	0.8%			Master's Degree	15.0%		
				Doctorate	1.7%		

Table 3: Participants demographics in Survey 2.

7.3 Survey 1: Participant-suggested Cookie Terms

Original Term	Alternate Terms Suggested
Strictly Necessary	Essential (+1), Required (+1), Secure (+1), System (+1), Active, Authentication, Basic, Choices, Cookies Necessary for Function, Cookies that Benefit You, Crucial, Functionality, History, Identification, Navigation, Needed for Functional Operation, Operational, Practical, Preferences, Support Cookies, User, User-friendly, Visits, Web Interface
Functionality	Personalized (+1), Algorithmic Preference Selection, Choices, Custom, Customization, Ease of use, History, Individualized, Memory, Necessary, Personal, Personalization, Personalized Preferences, Practical, Saved, User, User-Customized, User Enhancement, Website Customization, Your Preferences
Performance	Anonymous Information, Data, Error, Necessary, Page Visitation Data, Practical, Required, Tracking, Troubleshooting, User, User-Friendly, Website Functionality
Advertising/Targeting	Ad (+1), Personal Data Tracking (+1), Tracking (+1), External Advertising Information, Personal Advertisement Generating, Relevance, Solicitor, Surveillance cookies

Table 4: Participant-suggested cookie terms split out by categories defined by original cookie category terms. (+1) indicates that more than one participant suggested this term. Suggested terms were removed that were clearly irrelevant to the category (ex. eCommerce Cookies as an alternative for Strictly Necessary), irrelevant on the whole (ex. Cookie Monster), or biased (ex. Stalker Cookies).

7.4 Survey Consent Form

ONLINE CONSENT FORM FOR PARTICIPATION IN RESEARCH

Procedure:
The survey will take approximately 5 minutes to finish. This survey was created by Carnegie Mellon University researchers.

Purpose:
The purpose of this study is to research ways to describe different types of web cookies.

Participant Requirements:
Participants must meet the following requirements to participate in this study:

1. 18 years of age or older

2. Based in the United States

3. Can speak and understand basic English

Risks and Benefits:
There are no specific benefits to participants.
   The risks and discomfort associated with participation in this study are no greater than those ordinarily encountered in daily life or during other online activities.

Confidentiality:
The survey results will be stored using a unique identifier that does not reference your name, email address, or other personal information. We may share anonymized data with people outside of the research team. This may include sharing de-identified data and findings with other researchers, as well as publishing our findings in academic publications.

Voluntary Participation Section:
Your participation in this research is voluntary. You may stop the survey at any time. You may print a copy of this consent information for your records.

7.5 Survey Recruitment Text (Prolific)

Survey 1
The purpose of this Carnegie Mellon University (CMU) research study (STUDY2022_00000081) is to assess the comprehension

of various terms that describe types of web cookies. Participants must be 18 years of age or older and based in the United States. The survey is expected to take approximately 10 minutes to complete, and participants will be compensated $1.59 for their time. Participation in this study is voluntary, and you have the right to exit the survey at any point by closing your browser tab or window. Any questions or concerns can be directed to the study's principal investigator, Rachna Sasheendran, at rsasheen@andrew.cmu.edu. To participate, please accept the study in the Prolific platform and begin by reading our consent form. Thank you for your interest!

Survey 2
The purpose of this Carnegie Mellon University (CMU) research study (STUDY2022_00000081) is to assess the comprehension of various terms that describe web cookies. Participants must be 18 years of age or older and based in the United States. The survey is expected to take approximately 5 minutes to complete, and participants will be compensated $1.00 for their time. Participation in this study is voluntary, and you have the right to exit the survey at any point by closing your browser tab or window. Any questions or concerns can be directed to the study's principal investigator, Rachna Sasheendran, at rsasheen@andrew.cmu.edu. To participate, please accept the study in the Prolific platform and begin by reading our consent form. Thank you for your interest!

7.6 Survey 1 Questions

Consent Questions

1. I am age 18 or older

    (a) Yes

    (b) No

2. I have read and understood the consent form.

    (a) Yes

    (b) No

3. I want to participate in this research and continue with the task.

    (a) Yes

    (b) No

Demographic and Background Questions

4. What is your age group?

    (a) 18 - 24

    (b) 25 - 34

    (c) 35 - 44

    (d) 45 - 54

    (e) 55 - 64

    (f) 65 - 74

    (g) 75 - 84

    (h) 85 or older

5. Which is your gender?

    (a) Male

    (b) Female

    (c) Non-Binary

    (d) Prefer to self describe *(open box to describe)*

(e) Prefer not to say

6. What is the highest degree or level of schooling that you have completed?

   (a) Less than high school diploma

   (b) High school diploma or GED

   (c) Some college education but no degree

   (d) Associate's Degree

   (e) Bachelor's Degree

   (f) Master's Degree

   (g) Doctorate

   (h) Other *(open box to describe)*

7. On a scale from 1 (Extremely Unfamiliar) to 5 (Extremely Familiar), please rate your familiarity with privacy concepts (such as web cookies). *(options are displayed horizontally)*

   (a) 1 (Extremely Unfamiliar)

   (b) 2

   (c) 3

   (d) 4

   (e) 5 (Extremely Familiar)

   Cookie Category Terms Questions

8. These cookies are essential in order to enable you to move around the website and use its features, such as accessing secure areas of the website. Without these cookies services you have asked for, like shopping baskets or e-billing, cannot be provided.

   Please indicate how well you think each word/phrase below reflects this description. *each option can be rated on the following Likert scale: Not well at all, Slightly well, Moderately well, Very well, Extremely well*

   (a) Functional Cookies

   (b) Necessary Cookies

   (c) Status Cookies

   (d) Strictly Necessary Cookies

   (e) User State Cookies

   (f) Website Function Cookies

9. These cookies allow the website to remember choices you make (such as your user name, language or the region you are in) and provide enhanced, more personal features. For instance, a website may be able to provide you with local weather reports or traffic news by storing in a cookie the region in which you are currently located. These cookies can also be used to remember changes you have made to text size, fonts and other parts of webpages that you can customize. They may also be used to provide services you have asked for such as watching a video or commenting on a blog. The information these cookies collect may be anonymized and they cannot track your browsing activity on other websites.

   Please indicate how well you think each word/phrase below reflects this description. *each option can be rated on the following Likert scale: Not well at all, Slightly well, Moderately well, Very well, Extremely well*

   (a) Consistent Experience Cookies

   (b) Ease-of-access Cookies

   (c) Ease-of-navigation Cookies

   (d) Functional Cookies

(e) Functionality Cookies

(f) Personalized Experience Cookies

(g) Preferences Cookies

(h) Session Enhancing Cookies

10. These cookies collect information about how visitors use a website, for instance which pages visitors go to most often, and if they get error messages from web pages. These cookies don't collect information that identifies a visitor. All information these cookies collect is aggregated and therefore anonymous. It is only used to improve how a website works.

Please indicate how well you think each word/phrase below reflects this description. *each option can be rated on the following Likert scale: Not well at all, Slightly well, Moderately well, Very well, Extremely well*

(a) Analytics Cookies

(b) Interaction Statistic Cookies

(c) Metric Cookies

(d) Performance Cookies

(e) Site Statistic and Error Handling Cookies

(f) Statistics Cookies

(g) Visitation and Error Checking Cookies

11. These cookies are used to deliver advertisements more relevant to you and your interests. They are also used to limit the number of times you see an advertisement as well as help measure the effectiveness of the advertising campaigns. They are usually placed by advertising networks with the website operator's permission. They remember that you have visited a website and this information is shared with other organizations such as advertisers. Quite often targeting or advertising cookies will be linked to site functionality provided by the other organization.

Please indicate how well you think each word/phrase below reflects this description. *each option can be rated on the following Likert scale: Not well at all, Slightly well, Moderately well, Very well, Extremely well*

(a) Advertising Cookies

(b) Individual Behavioral Advertising Cookies

(c) Marketing Cookies

(d) Targeting Cookies

(e) Third Party Advertising Cookies

(f) Personalized Advertising Cookies

Open-Ended Questions *After each question in questions 8-11, we asked the following two open-ended questions:*

12. Why do you think these terms work particularly well or not well? *(open answer box)*

13. Do you have any other suggestions for words/phrases that could reflect this description well? If writing multiple suggestions, please separate them using a comma. *(open answer box)*

Prolific Completion Validation Question

14. What is your Prolific ID? *(open answer box)*

7.7 Survey 2 Questions

Consent Questions

1. I am age 18 or older

(a) Yes

(b) No

2. I have read and understood the consent form.

    (a) Yes

    (b) No

3. I want to participate in this research and continue with the task.

    (a) Yes

    (b) No

Demographic and Background Questions

4. What is your age group?

    (a) 18 - 24

    (b) 25 - 34

    (c) 35 - 44

    (d) 45 - 54

    (e) 55 - 64

    (f) 65 - 74

    (g) 75 - 84

    (h) 85 or older

5. Which is your gender?

    (a) Male

    (b) Female

    (c) Non-Binary

    (d) Prefer to self describe *(open box to describe)*

    (e) Prefer not to say

6. What is the highest degree or level of schooling that you have completed?

    (a) Less than high school diploma

    (b) High school diploma or GED

    (c) Some college education but no degree

    (d) Associate's Degree

    (e) Bachelor's Degree

    (f) Master's Degree

    (g) Doctorate

    (h) Other *(open box to describe)*

7. On a scale from 1 (Extremely Unfamiliar) to 5 (Extremely Familiar), please rate your familiarity with privacy concepts (such as web cookies). *(options are displayed horizontally)*

    (a) 1 (Extremely Unfamiliar)

    (b) 2

    (c) 3

    (d) 4

    (e) 5 (Extremely Familiar)

Cookie Comprehension Questions *The order of these questions was randomized for every survey participant.*

8. What are _____ cookies? *(Asked once each for the following terms: Functional, Necessary, Strictly Necessary, and Website Function)*

    (a) Cookies that are needed for the website to work properly

    (b) Cookies that are needed for collecting certain metrics

    (c) Cookies that are needed for determining your location

    (d) I don't know

9. What are _____ cookies? *(Asked once each for the following terms: Functionality, Personalized Experience, Preferences)*

    (a) Cookies that are needed for the website to work properly

    (b) Cookies that help personalize the website's services for you

    (c) Cookies that are given lower priority than other cookies on the website

    (d) I don't know

10. What are _____ cookies? *(Asked once each for the following terms: Performance, Analytics)*

    (a) Cookies that help measure and improve website features

    (b) Cookies that are given priority over other cookies on the website

    (c) Cookies that make the website run faster

    (d) I don't know

11. What are _____ cookies? *(Asked once each for the following terms: Targeting, Advertising, Marketing, Third Party Advertising, Personalized Advertising)*

    (a) Cookies that are used for delivering personalized advertisements

    (b) Cookies that help users navigate the website

    (c) Cookies that are needed for determining your location

    (d) I don't know

Prolific Completion Validation Question

12. What is your Prolific ID? *(open answer box)*

7.8 Code-book

All categories	Codes	Sub section	Acronym used in our coding	Definitions	Key terms
Positive: "why the terms work well"	**Intuition**		**PI**	Users intuitively feel like this makes sense	"I think", "It seems", "I feel" "laymen"
	Comprehension		**PCOMP**	Users understand how the word matches the definition	"understand", "describes", "summarizes"
	Concise		**PCON**	Users find the terms to be short and straightforward	"short", "to the point", "quick" "simple"
Negative: "why the terms don't work well"	**Confusing**		**NC**	Users are confused about how the term matches the definition	"confused" "not sure"
		Vague	*NCV*	Users are confused and they think the term isn't specific enough and it's too general	"general", "broad", "not specific enough"
		Conflicting	*NCC*	Users are confused and they think that the term is in conflict with the definition provided.	"does not match", "inaccurate"
	Wordy		**NW**	Users found that the terms were too wordy or lengthy to comprehend them properly	"too long" "wordy" "verbose" "too many words"
Other	**Unclear**		**OU**	Not enough information to draw conclusions	Not Applicable

Figure 1: Code-book for Qualitative analysis of free-responses from Survey 1

7.9 Frequency of responses based on code-book

	Category 1: Strictly necessary	Category 2: Functional Cookies	Category 3: Performance cookies	Category 4: Targeting cookies
PCOMP	58	46	42	47
PCON	4	6	7	10
PI	12	24	18	17
NCC	4	2	3	5
NCV	14	10	15	6
NW	0	4	9	4
NC	1	0	0	0
OU	8	6	7	7

Figure 2: Frequency of responses for each code section and sub-section in each cookie category from Survey 1 as per the above Code-book