Generational Differences in Understandings of Privacy Terminology

Charlotte Moremen
Pomona College
cbma2019@mymail.pomona.edu

Jordan Hoogsteden*
Harvard Law School
jhoogsteden@jd26.law.harvard.edu

Eleanor Birrell
Pomona College
eleanor.birrell@pomona.edu

ABSTRACT

Prior work has consistently found that people have miscomprehensions and misunderstandings about technical terms. However, that work has exclusively studied general populations, usually recruited online. This work investigates the relationship between generational cohorts and their understandings of privacy terms, specifically cohorts of elementary school children (aged 10-11), young adults (aged 18-23), and retired adults (aged 73-92), all recruited offline. We surveyed participants about their understanding of and confidence with technical terms that commonly appear in privacy policies. We then moderated a post-survey focus group with each generational cohort in which participants discussed their reactions to the actual definitions along with their experience with technical privacy terms. We found that young adults had better understandings of technical terms than the other generations, despite all generations reporting being regular Internet users. Participants across all generational cohorts discussed themes of confusion and frustration with technical terms, and older adults particularly reported a sense of being left behind. Our results reinforce the need for improvement in the presentation of information about data use practices. Our results also demonstrate the need for more focused research and attention on the youngest and oldest members of society and their use of the Internet and technology.

KEYWORDS

privacy, technical terms, user study

1 INTRODUCTION

In modern online life, privacy policies are everywhere, and people are frequently asked—or presumed—to consent to data practices described by a privacy policy. For instance, in a 2019 Pew Research Center study, 25% of adults said they were asked to agree to the terms and conditions of a company's privacy policy on an almost daily basis, while 32% said this happens about once a week [6]. Despite the central role these policies play in current models of privacy, these documents are lengthy [22, 51], require an advanced reading level [22], and contain ambiguous language [10]. Part of privacy policies' inaccessibility comes from the technical terms they use. Many technical terms that commonly appear in privacy policies are either not understood or are misinterpreted by many people [80]. Consequently, most Internet users either skim over privacy policies or do not read them thoroughly enough to understand their intended meaning [7, 59, 69].

Although existing critiques of privacy policies are compelling, research that looks generally at the broader public—particularly research conducted through online platforms such as Prolific and Amazon Mechanical Turk—likely underestimates the scope of the problem. That prior work systematically excludes two populations that are generally less technically-sophisticated than the overall population: children and older adults.

Most research—and all current online crowdsourcing platforms—explicitly excludes minors under age 18. However, children are active Internet users. In 2016, most children had access to social media by age 12, with 23% of children aged 8-12 also having their own social media accounts [43]. Internet usage by children has increased since then [57, 71], and 97% of American children aged 3-18 have home Internet access [58]. While collection of data about children under 13 generally requires parental consent, some regulations also require that data practices be transparent to children. For example, Recital 58 of the E.U.'s General Data Protection Regulation (GDPR) states, "Given that children merit specific protection, any information and communication, where processing is addressed to a child, should be in such a clear and plain language that the child can easily understand" [21]. The Transparency Guidelines issued by the Article 29 Working Party further clarify that children do not lose their right to transparency just because consent has been given by a parent [5]. Beyond these legal requirements, prior work has shown [16, 47]—and our results confirm—that many children use these services without parental consent, reinforcing the importance of understanding how children understand or misunderstand these technical terms in privacy policies.

Older adults are also active Internet users. In 2022, 75% of adults aged 65 and older used the Internet, and 45% of that age cohort used social media [24]. While some older adults participate in studies on crowd-sourcing platforms, that age group is underrepresented, and the older online participants who are on these platforms are not representative of their generation. Older adults who complete tasks on crowdsourcing platforms are significantly more technically sophisticated than their age group overall, which results in significantly different responses between online survey platform users compared to general older adults who use the Internet [68, 79]. It is therefore valuable to understand how the general population of older adults understand or misunderstand technical terms.

This work seeks to address this gap in current research by exploring generational differences in understandings of technical terms that appear in privacy policies. Our goals are (1) to understand how comprehension of terms varies between age groups and (2) to explore what gets missed by online study populations. To do

*Work was conducted while Hoogsteden was at Pomona College.

Proceedings on Privacy Enhancing Technologies 2024(3), 589–605

https://doi.org/10.56553/popets-2024-0094

this, we conducted in-person user studies with three age cohorts: children aged 10-11, young adults aged 18-23, and older adults aged 73-92. These groups were then classified by generational cohort: children aged 10-11 are members of "Generation Alpha", young adults aged 18-23 are members of "Generation Z", and retired adults aged 73-92 are a mix of "Baby Boomers" and "The Silent Generation". Since there were too few Baby Boomers to evaluate separately, we combined Baby Boomers and the Silent Generation into a single Retirement Generation or "Generation R". All participants were recruited offline through organizations in our community (an elementary school, a consortium of undergraduate colleges, and a retirement community). With each age cohort, we surveyed participants about their interpretation of eight common terms found in privacy policies and about their confidence in their definitions. We then conducted post-survey discussions to further explore each generation's opinions about and experiences with privacy policies, privacy terms, online privacy, and the Internet.

We found that all generational groups we surveyed are active online and have encountered privacy policies. Regardless of age, participants expressed feelings of frustration and confusion with these policies and terminology, and there were some terms that all age groups struggled to define (e.g., "pixel tag"). However, we also found differences between age groups. Both Generation Alpha and Generation R correctly defined fewer terms than Generation Z, with differences for most of the individual terms as well. Generation R was more likely to report low levels of confidence in their understanding of privacy terms. They also described feeling "ignorant", "stupid", and "really old" when faced with privacy terms.

Our results confirm that all age groups—including children and senior citizens—actively engage online, and that meaningful understanding of privacy policies across all ages is critical to establish informed consent. Our results also indicate that users across all age groups do not fully understand how their data is collected and used when they agree to privacy policies, and that misunderstandings and confusions about technical terms that appear in privacy policies preclude informed consent. Moreover, our results show that in the realm of online data and privacy, there are differences in behavior, attitudes, and understanding between different generations. These results suggest that future usable privacy research needs to include children and older adults—especially those who are not enrolled in online survey platforms—or needs to be replicated with these underrepresented populations in order to enable equitable privacy for people of all ages.

2 RELATED WORK

This research extends both prior work relating to comprehension of privacy policies and prior work exploring the privacy needs of specific age cohorts. However, to the best of our knowledge, this is the first work to combine these two lines of work by studying how children and older adults understand technical terms that appear in privacy policies.

Understandings of Privacy Terms. Much work had been conducted on privacy policies generally [1, 2, 4, 14, 32, 41, 46, 48, 49, 55, 88, 96], and a few papers have looked specifically at how users understand privacy terms.

Tang et al. [80] examined users' understanding of technical terms that appear in privacy policies and investigated how misunderstandings impact comfort with described data practices. They found that confusion and misunderstandings were common. On average, people answered just 40% of answers correctly. For 15 of the 22 terms studied, less than half of participants were able to correctly define the term on a multiple-choice question. However, their study recruited participants on Amazon Mechanical Turk, so it excluded children under 18. Moreover, online studies are known to not be representative of older adults [68, 79]. Our work extends their results by identifying differences between different age cohorts, including children and older adults recruited offline, and by collecting and analyzing rich, qualitative data.

Other work looking at how people interpret privacy-related terms in particular contexts also found high levels of misunderstandings. User studies have consistently found that users misunderstand the technical term "privacy policy". A 2005 study about the online shopping behavior of American consumers found that 75% falsely believed that the presence of a privacy policy meant that a website would not share their information with other websites or companies [81]. More recently, a 2014 survey [74] and a subsequent longitudinal study [82] found that a majority of Internet users continue to hold this misconception. Work looking at understandings of mobile app permissions found that users answered 21% of permission comprehension questions correctly, and just 2.6% of respondents answered all three questions correctly [25]. Many users could not connect the resource-specific technical terms used in permission names to particular risks that would be enabled by those permissions. In an online study that examined the impact of design elements on cookie consent decisions, less than half of participants could correctly define "performance cookies", and only 16% could define "functional cookies" [31]. Recent work exploring the usability of app privacy labels found that some terms in those labels, e.g., "Data Used to Track You" were commonly misinterpreted [45, 97].

Comprehensibility of Privacy Policies. Misunderstandings and confusions about the meanings of technical terms are one piece of a larger body of work evaluating and critiquing the comprehensibility of privacy policies. Several independent projects have analyzed the text of privacy policies to quantify readability using standardized metrics such as the Flesch Reading Ease Test [26]. These studies have consistently found that privacy policies are difficult to comprehend and are often written at a level that surpasses the educational levels of many of the people they are intended to inform [3, 8, 9, 13, 20, 22, 28, 34, 35, 39, 44, 52, 53, 72, 73]. Studies that looked at healthcare privacy policies found that none of the policies examined were readable by a majority of English speaking Americans [28], and that on average 80% of people living in areas surrounding the hospitals whose privacy policies were studied were not at the reading level required by these policies [8]. Additional work has studied the readability of financial privacy policies [3, 34, 44]. Most comprehensively, Fabian et al. [22] analyzed the privacy policies of 50,000 English-speaking websites, finding that these policies were difficult to comprehend and required college-level reading levels. Moreover, privacy policies frequently employ linguistic techniques such as euphemisms, passive

verbs, and modality markers ("may") that can obscure the meaning of the policy [64, 65]. Poor writing style, uncommon words, and difficult-to-read formatting are also common [34].

Additionally, user studies have consistently confirmed that privacy policies are difficult for people to understand [75, 81]. Proctor et al. [66] found that people were only able to answer approximately 50% of comprehension questions about privacy policy practices despite these policies being written at their education level. Vu et al. [86] found that participants at the reading level required by the policies displayed poor overall understanding of their contents [86]. Singh et al. [73] evaluated the readability of privacy policy statements from ten popular websites and found that users do not completely understand the contents of privacy policies. For half of the websites, no participant passed the target comprehension threshold (a Cloze score above 0.6).

Children and Privacy. To the best of our knowledge, no prior work has looked at children's understandings of technical terms that appear in privacy policies. However, children under 18 have occasionally been included in studies relating to privacy, and prior work suggests that children's attitudes can differ from most adults. A 2019 interview study with 11 children aged 7-10 found that when considering potential harms from technology, children were the most fearful of physical harm (like being attacked by a robot) and loss of attachment (being taken away from their families) [95]. In 2021, Hiniker et al. [33] found that children absorb linguistic patterns from conversational agents (such as Siri or Alexa) and apply them to conversations with people, despite never being told about the pattern by a human. However, they did not investigate children's understandings of privacy and security in relation to conversational agents.

Miltgen et al. [54] examined 139 online users in European countries and found that 18.5% of participants aged 15-24 said they were unconcerned about privacy, while only 8.1% of those aged 25-70 agreed. The younger group felt more optimistic about their personal data and confident in their ability to prevent data theft or misuse. However, their work did not include participants under the age of 14. Kumar et al. [40] conducted an interview study with children under 12 and found that children demonstrated some grasp of privacy issues—e.g., identifying some data as sensitive and being comfortable sharing information with certain people and not others—but that they sometimes failed to recognize privacy issues. Zhang-Kennedy et al. [98] conducted interviews with parent-child pairs in which they found that children's threat models differed significantly from parents' threat models, with internal threats from family members having higher saliency for children.

To investigate children's understandings of online privacy and security, Zhao et al. [99] conducted focus groups with children aged 6-10 in the UK. The focus groups walked the children through hypothetical scenarios about online privacy using a cartoon koala bear named Bertie who likes playing on her iPad and who experiences situations such as implicit video promotions (for example, algorithmic queuing in YouTube), in-app pop-ups, and downloading "free" games (most of which remain free by collecting user data). They found that while children were cautious of certain features such as popups, they were unaware of more insidious uses of their data, such as algorithmic queuing and "free" apps.

Another focus group conducted with children aged 7-13 in 2021 investigated how children would design online privacy warnings for their peers [15]. Only 37/150 drawings had the 3 attributes of an effective warning message (Attention, Knowledge, and Compliance). 60 drawings used words to communicate risk, while only 13 used pictorial representations. These results indicate that children struggle to fully comprehend the risks they are undertaking when going online.

Given the findings that children are largely unaware of and uneducated about online privacy and security, a natural question arises: have digital citizenship programs had any impact on children's knowledge of online security and privacy? In 2023, Jones et al. [37] investigated the effectiveness of Google's "Be Internet Awesome" (BIA) program, designed to educate children about privacy online. American elementary and middle schools students who completed the BIA curriculum were significantly more likely to say they check privacy settings and have their social media on partly private or completely private. They also understood terms including "catfishing", "hacker", and "trolling" significantly better than students in the control group. However, neither group consistently understood technical terms will, with only 31.6% of students who completed BIA correctly defining "digital footprint".

Legal regulations also impact children's privacy online. However, legal protections for children online vary by state, country, and region of the world, and they do not fully protect children's privacy. For example, Kuznicka-Błaszkowska [42] highlighted the issue of parents posting embarrassing "funny" stories, photos, or videos about their children online that the children may not have consented to or did not have the cognitive ability to truly consent to. Recital 65 of the GDPR [21] emphasizes that children (and former children) have a right to delete personal data, but it cannot fundamentally go back and undo the embarrassment the child experienced or the memories of the people who witnessed it when it was online. Kuznicka-Błaszkowska also points out flaws in the Children's Online Privacy Protection Act (COPPA) [83], the federal law protecting Internet privacy for children in the United States. For example, the fact that it only encompasses the acts of private companies.

While both education and legal protections are valuable tools for increasing children's online security, their current implementations are imperfect. Additionally, minors' disregard and ignorance of privacy concerns sometimes puts them at risk. For example, many minors lie about their age online to gain access to social networks when creating their accounts [16]. This behavior inadvertently circumvents the precautions that social media platforms put in place to protect minors.

Older Adults and Privacy. Although most large-scale user studies include some older adults, the number of older adults who participate in such studies is frequently small, and older adults on crowd-sourcing platforms are not representative of that age group overall [68, 80]. However, older adults—and the differences between this population and younger adults—have been the focus of some prior privacy studies.

In a focus group, researchers found that members of the Silent Generation had a general fear of the Internet and what it meant for their privacy [36]. In contrast, Baby Boomers had more specific

concerns like hacking. At the same time, several felt more annoyed than threatened by online hacking and phishing attempts, assuming that their bank or computer would resolve the issue. A study conducted in India in 2021 interviewed multi-generational households and found that there is commonly a "technology manager", an adult in the house who oversees children and older adults' use of the Internet [56]. However, these technology managers can find it difficult to explain cybersecurity concepts to older adults as they may struggle to understand the concepts themselves.

In 2018, Elueze and Quan-Haase [19] conducted 40 interviews with older adults (65+) in Canada about their privacy concerns. They expanded on Westin's typology on privacy attitudes [90] and categorized participants into 5 labels: fundamentalist, intense pragmatist, relaxed pragmatist, marginally concerned, and cynical expert. 57% of respondents were pragmatists (42% were relaxed pragmatists and 15% were intense pragmatists), 25% were marginally concerned, 13% were fundamentalists, and 5% were cynical experts. Participants across groups shared concerns about surveillance, scams, and identity theft. The marginally concerned group had the least amount of concerns, and the cynical experts had several concerns.

Frik et al. [27] interviewed older adults in San Francisco about their perceptions of online privacy and security and thematically coded the interview transcripts. They found that 28 of 46 participants mentioned concerns about data collection and its lack of transparency, and several were concerned about their data being sold for profit. Additionally, several participants were concerned that older adults were specifically targeted for attacks because of their perceived vulnerability, while several did not think that attackers viewed them as "major consumers".

Ray et al. [67] conducted a similar study published the same year, interviewing 20 adults aged 60+ about their mental models for the term "privacy", both online and offline. Participants drew their mental models of privacy, which the research team later analyzed and thematically coded. They found that 18/20 drawings depicted feelings of fear or anger towards privacy invasion, with 12/20 depicting restricting use of technology to prevent privacy invasion. They suggested that improved transparency and a user-friendly approach could ease some of the privacy concerns of older adults. They additionally recommended privacy training for older adults about the benefits and importance of good privacy practices.

Overall, studies on older adult's online security and privacy understandings have found that while many older adults are concerned about their online privacy, they do not have the knowledge or tools to improve their security habits and are often left confused and fearful of the Internet.

3 METHODOLOGY

The goal of this work was two-fold: (1) to understand how understandings of technical terms that appear in privacy policies vary between generational cohorts and (2) to explore what gets missed by online study populations for research in this domain. To do so, we conducted in-person user studies with three generational groups: children aged 10-11, young adults aged 18-23, and older adults aged 73-92. With each age cohort, we qualitatively surveyed participants about their interpretation of eight common terms found in privacy policies as well as their confidence in their responses. We then conducted post-survey discussions with each age cohort to further explore each generation's opinions about and experiences with privacy policies, privacy terms, online privacy, and the Internet.

3.1 Subject Recruitment

We included four generations in our study: Generation Alpha, Generation Z, Baby Boomers, and the Silent Generation. Generational lines are often disagreed upon, especially for the youngest generations. For the purposes of this study, we used the Pew Research Center's definition of the Silent Generation as those born between 1928-1945 and Baby Boomers as those born between 1946-1964 [17]. The Pew Research Center defines Generation Z as those born in or after 1997, but they have not yet determined a chronological endpoint. Mark McCrindle and Ashley Fell, who coined the term "Generation Alpha" for the youngest people alive today, define the generation as starting in 2010 [50]. In this work, we define Generation Z as people born between 1997-2010 and Generation Alpha as those born after 2010.

We recruited a convenience sample for each generational group in October and November of 2022. For Generation Alpha, we contacted a local elementary school and asked to survey one of their older classes. After sending out consent forms to the parents, we met with 14 fifth graders (aged 10-11). The Generation Z group was recruited via flyers posted on a college campus. Participants scanned a QR code on the poster and filled out a Google Form to sign up, yielding 15 participants (aged 18-23). Members of the Silent Generation and Baby Boomers were recruited after we contacted a local retirement community and asked to visit. There were 11 participants from the retirement community: 8 members of the Silent Generation and 3 Baby Boomers. Since there were too few Baby Boomers to evaluate separately, all retirees were combined into Generation R. Most Generation R participants ranged from 73 to 92 years of age (birth years 1930-1949), although there was one outlier who was 61 years old (born 1961). All participants were residents of the United States. Detailed demographics of each generational cohort are summarized in Table 1.

		Gen A	Gen Z	Gen R
Birth year	2010-	14	0	0
	1997-2009	0	15	0
	1981-1996	0	0	0
	1965-1980	0	0	0
	1946-1964	0	0	3
	1928-1945	0	0	8
Gender	Woman/Girl	6	10	9
	Man/Boy	8	2	2
	Non-binary	0	2	0
	Self-describe	0	1	0
Race and Ethnicity	White-Eur.	5	12	11
	East Asian	0	2	0
	Black	1	1	0
	Latino	3	0	0
	Mixed	5	0	0

Table 1: Demographics of our three generational cohorts.

3.2 Study Protocol

We conducted our user study in three sessions, one for each generational cohort. Each session—which was conducted on-site at the elementary school, college, and retirement community respectively—was comprised of (1) a survey about people's understanding of eight common privacy terms and their confidence in their answers and (2) a post-survey focus group discussion that further explored each generation's opinions about and experiences with privacy policies, privacy terms, online privacy, and the Internet. The total time for each session, including survey, focus group discussion, introductions, transitions, and wrapping up was approximately 40 minutes.

Term Selection. To select our terms, we started with a list of 22 technical privacy terms that commonly appear in privacy policies constructed by Tang et al. [80]. To keep the length of our study—which included both a survey and a discussion—reasonable, we needed a shorter list of terms. We therefore cross-referenced the Tang list with terminology we found in the privacy policies of the most popular U.S. phone (Apple) [38], search engine (Google) [77] and social network (Meta) [78] in October 2022, and we eliminated terms that did not appear in any of those three privacy policies. We chose those commercial sectors for their broad appeal across all three age groups—children are unlikely to use Microsoft products, for example, as it is primarily a workplace technology. We then eliminated terms that are commonly understood, have no common misconceptions, or have no standard definition, as well as merging similar terms ("local storage" and "browser storage" collapsed into "cache", "session cookie" and "persistent cookie" into "cookie"). Finally, we added the term "biometric data"—which did not appear in Tang et al.'s study—to further explore established misconceptions of device fingerprinting as fingerprint-based identification. Our final term list therefore included the following eight terms:

(1) Privacy policy
(2) Cookie
(3) Pixel tag
(4) Cache
(5) Metadata
(6) Device fingerprinting
(7) Encryption
(8) Biometric data

Participant's definitions of terms were compared to correct definitions of the terms taken from the Computer Security Resource Center [12] or Wikipedia [91–94] in October 2022. These correct definitions are reproduced in Appendix B.

User Survey. Each generation's session began with a 15-minute survey. Two authors remained in the room with participants during the survey phase to answer any clarifying questions that arose (such as "how do I go to the next page?"). Participants chose whether to complete the survey online or on paper copies we provided. Participants were not permitted to search for definitions or ask one another or the research team for them. If participants attempted to ask the research team for the definition, they were encouraged to put their best guess or "I don't know" as their response.

The survey asked participants to define each term on our final term list with an open-ended response and to rate their confidence in their answers on a 5-point Likert scale from "very unconfident" to "very confident". In order to provide context, we provided an excerpt from the October 2022 version of the Google, Facebook, or Apple privacy policy that used each term. If only one of the three policies used the exact term, we used an except from that policy. If multiple of the privacy policies used the term, we used the excerpt that provided the best contextual information. Complete text of each selected excerpt is included in the survey protocol reproduced in Appendix A.

After the term definition section, we asked participants multiple-choice Likert-scale questions about their overall confidence in their responses and their comfort with technology, as well as if the survey prompted any interest in learning more about privacy terms and concepts. We also asked participants for basic demographic information (birth year, gender, and race) but kept responses anonymous. We ended with an open-ended question: "Lastly, if any, what do you think the consequences are of not knowing what these terms mean?" The survey was proofed for readability at a fifth-grade level by an author with experience in childcare and child education. A copy of the full survey protocol is provided in Appendix A.

Focus Group Discussions. After all participants finished the survey, we conducted a 15 minute focus group discussion with each cohort. Open discussion was encouraged, with prompting questions such as:

- Does anyone have any comments or questions they'd like to start with?
- Was there anything about the experience you'd like to talk about that wasn't mentioned in the survey?
- How did you feel taking the survey? Why? (e.g., annoyed, bored, stressed, excited etc.)
- Before taking this survey, how much attention did you give to privacy policies?
- What about individual privacy? (e.g., passwords, personal devices)
- Was there a term you found particularly challenging?

Given ethical concerns about working with vulnerable populations, we agreed with our IRB not to record audio during our focus groups. One author anonymously transcribed quotes from each session while another author led the focus group.

3.3 Data Analysis

Our user study resulted in three types of data: open-ended qualitative survey responses, quantitative survey responses, and notes from focus-group discussions. In this section, we describe how we analyzed these three classes of data.

3.3.1 Qualitative Survey Responses. The core of our data resulting from this user study was a rich set of qualitative responses about how people in different generations understand various technical terms relating to privacy. We also had qualitative survey data about what people thought the implications of not understanding technical terms might be. We performed two types of analysis on these open-ended survey responses: (1) deductive coding for correctness and (2) thematic coding to identify patterns in how participants understood those terms.

Term	Fully Understands	Big Idea	Vague Idea	Doesn't Know	Miscomprehension
Metadata	*data that describes/ identifies/categorizes other data*	*information about when, where, and on what device you took your pictures on.*	*data involving personal information to get a wider scope of data, like location etc*	*idk*	*large amount of data*
Device fingerprinting	*anything that can be used to identify a particular device, like IP address or something*	*things that are able to locate your specific device and make it targetable to other people*	*how each device is different*	*don't know?*	*the fingerprints you put onto devices via touch ID.*

Table 2: Examples of how qualitative responses about technical terms were coded for correctness.

Coding for Correctness. We deductively coded participants' free-response definitions of each technical term using a five-point scale:

(1) Subject fully understands concept
(2) Subject understands the big idea, but not specifics
(3) Subject has a vague idea
(4) Subject doesn't know
(5) Subject has incorrect understanding

For example, a definition of metadata was coded as "fully understands" if it included a general characterization of data about other data. Responses that gave particular examples (e.g., when and where a photo was taken) were coded as understanding the big idea, and responses that mention data in a way consistent with metadata were coded as understanding the vague idea. Answers that explicitly acknowledged that the subject did not know the term were coded "doesn't know", and responses that were inconsistent with the correct definition (e.g., a response about metadata talking about scale of data or talking about the metaverse) were codes as incorrect. Examples of how specific responses were coded for correctness are presented in Table 2.

We double coded all responses for correctness. For our first round of coding, we compared responses and resolved disputes after each term for one generational group. After that, disagreements on codings were discussed and resolved after both researchers had individually finished their codes for the remaining two generational groups. All three of the generational codings yielded a Cohen's kappa coefficient above 0.6, indicating reasonable agreement. Generation Alpha's Cohen's Kappa was 0.8176, Generation Z's Cohen's Kappa was 0.6232, and Generation R's Cohen's Kappa was 0.7005. Generation Z's lower Cohen's Kappa is due to the fact that Generation Z's responses were the most verbose and the most likely to include multiple different ideas. Some responses from this generational cohort started as showing an understanding of the term but then deviated (or vice versa), leading us to reconcile our code assignments retroactively.

Thematic Coding. After familiarizing ourselves with our dataset, we thematically coded qualitative responses using a three-step process: (1) two researchers independently performed open-coding on the full set of responses, (2) we met as a group to discuss the set of open codes and group them into themes for each question, and finally (3) we jointly deductively coded the full set of responses using our set of identified themes.

3.3.2 Quantitative Survey Responses. In addition to the qualitative questions, our survey also included quantitative questions about confidence levels, comfort with technology, and personal demographics. Since our small sample size precluded statistical analysis, we focused on providing descriptive statistics and observations about differences observed in how different generational groups responded to these quantitative questions.

3.3.3 Focus Group Data. Given the inclusion of vulnerable populations, we elected not to record focus group discussions. As a result, the data collected during these focus groups is primarily comprised of individual quotes. Additionally, we polled each focus group about various topics and noted the responses. We provide descriptive statistics about polling results and include representative quotes from our focus groups to augment the analysis of our survey data.

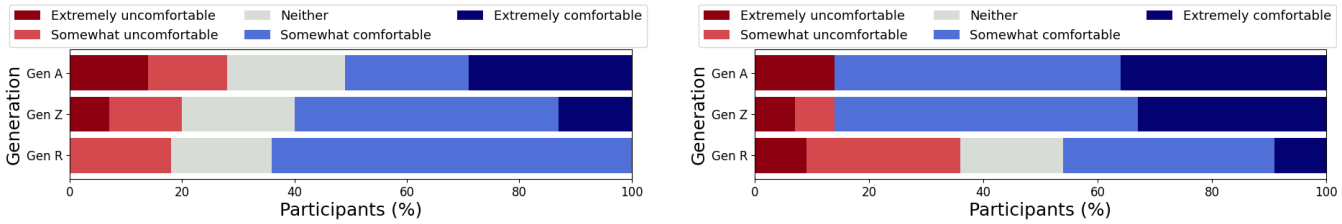
3.4 Ethical Principles

We took steps to ensure that informed consent was granted in advance for all participants in our study. For participants in Generation Alpha (who were minors under 18), we obtained signed parental consent prior to conducting the study. The children were additionally informed that their participation in the survey was voluntary and they did not have to respond to anything they did not wish to. All children assented to participate. For participants in Generation Z and Generation R, we obtained signed consent from each participant. For Generation R, we also confirmed with the retirement community staff that all participants were capable of granting informed consent.

We were careful to ensure that no participant data could be linked to individual participants. Consent forms were handled separately from surveys. All survey data was anonymous and contained no identifying information. To avoid recording any identifying information, we elected not to record our focus group discussions. Instead, we manually took notes and transcribed notable quotes.

To compensate people for their time, participants in Generation Z and Generation R received a $10 gift card to local businesses. Generation Alpha participants were compensated with a fidget toy[1].

[1]Since children may perceive small payments as significant financial sums [89], we believe toys are less likely than payments to exert undue influence on children's assent decisions [11].



(a) Computers **(b) Smartphones and Tablets**

Figure 1: Reported level of comfort with technology

Due to the inclusion of children under 18, this study was granted a Full Board Review and was approved in advance by the Institutional Review Board at our institution.

3.5 Limitations

This work, by design, includes age cohorts that are rarely studied by the usable privacy and security community. However, systematic biases in other demographics might limit the generalizability of our results. 25/40 of our participants from the three generational groups identified as women, and 28/40 were of White European descent. Furthermore, although not a demographic category captured by our survey, our data may also be socioeconomically limited. The colleges we conducted our study at are private liberal arts schools, the elementary school is located in a community where the cost of living is 57% higher than the national average [63] and where all students use tablets to complete schoolwork, and the retirement community vets residents based on their lifelong philanthropic work. All three of these environments tend to draw more wealthy people. It is therefore possible that our data is missing aspects of the national experience, perspective, and knowledge of technical terminology.

Future research should study larger sample sizes from different geographical communities and samples that are representative of the overall population. Nonetheless, we believe this exploratory work is valuable to the community by identifying important new directions from small-scale qualitative data.

4 RESULTS

We analyzed our data to understand general trends and generational differences in experience, understanding, confidence, and attitude towards privacy terms. We also evaluated how our results compare to prior online studies.

4.1 Experience With Technology and Privacy

We asked participant to rate how comfortable they were with computers and with smartphones or tables. Most participants described themselves as comfortable with both technologies. However, we observed some differences between different generations. The younger cohorts were very comfortable with smartphones and tablets, with 12/14 children in Generation Alpha and 13/15 Generation Z participants saying they were somewhat or very comfortable with these devices. This is consistent with the fact that usage of these devices was high among all participants from both of those cohorts—all Generation Alpha participants had school-issued iPads, and all Generation Z participants had and regularly used smartphones. By contrast, only 5/11 Generation R participants said they were comfortable with smartphones and tablets. We also observed different trends for comfort levels with computers. Although none of our retired cohort said they were very comfortable with computers, only two described themselves as uncomfortable with computers (most said they were "somewhat comfortable"). Generation Alpha was the most likely to self-identify as uncomfortable with computers, suggesting that experience with phones and tablets does not generalize to general comfort with technology for children of this age. These results are shown in Figure 1.

During the focus groups, we asked about prior experience with privacy policies. All participants in all three generational groups had seen a privacy policy before, and a majority of participants in each generational cohort (9/13[2] Generation Alpha participants, 11/15 Generation Z participants, and 8/11 Generation R participants) reported that they had seen a privacy policy within the last month. While these numbers are slightly lower than the 81% of Americans who reported being asked to agree to a privacy policy in the last month in a prior survey [6], these results suggests that privacy policies—and their contents and terminology—are relatively prevalent in people's online lives for Internet users of all ages. Unsurprisingly, many participants also admitted to not reading privacy policies before using a website or online service. For example, all 15 Generation Z participants responded that they pay attention to privacy policies "none of the time". However, 8/11 Generation R participants, 10/15 Generation Z participants, and 9/13 Generation Alpha participants said they would be interested in learning more about online privacy and security topics.

Generation Alpha subjects expressed notably higher levels of suspicion towards sharing personal information compared to the two adult cohorts. All 13 children who participated in the focus group discussion reported that they regularly make up personal data when asked for it, confirming prior findings that children frequently lie to appear older on the Internet [16]. Some participants even cited specific streaming platforms and gaming websites that they lie about their age to use: one of our participants remarked that "Hulu thinks I'm 21". When probed further on the topic of providing personal information, one of the Generation Alpha participants commented, "if it asks for personal info it's sus".

[2]One of the 14 Generation Alpha participants excused themself to go to the bathroom and was not part of the focus group discussion.

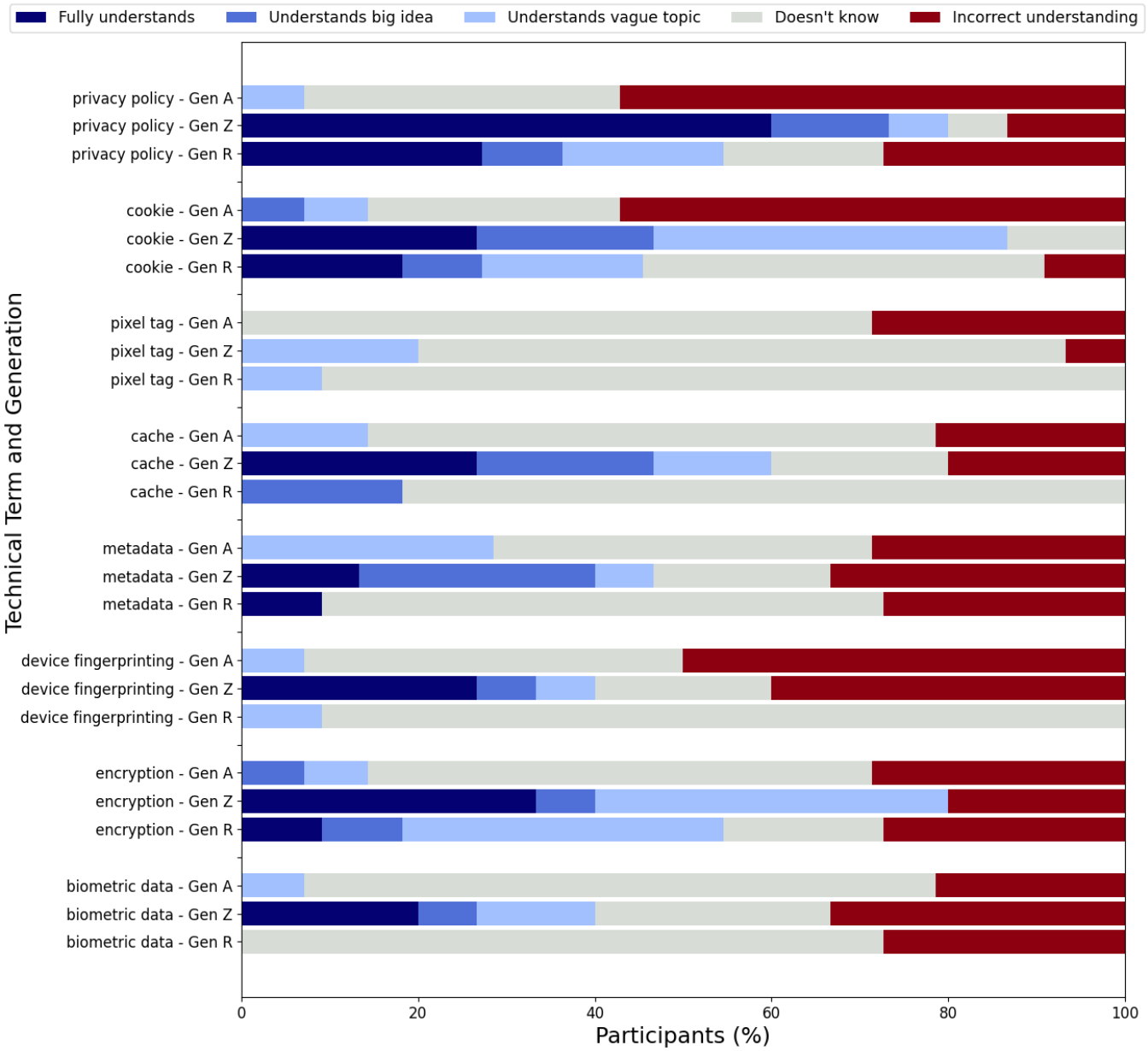


Figure 2: How well people understood privacy terms across different generations (qualitative coding).

4.2 Understandings of Privacy Terms

To understand how participants in each generational cohort understood our eight technical privacy terms, we both deductively coded responses for correctness and thematically coded responses as described in Section 3.3. This subsection describes the results of both analyses. The results of the correctness coding are also depicted in Figure 2.

Privacy Policy. Consistent with prior work [80], "privacy policy" was the most broadly understood term included in our study. However, we observed distinct generational differences in how well people understood this term. While 9/15 Generation Z participants gave fully correct definitions, only 3/11 Generation R participants and 0/14 Generation Alpha participants gave a fully correct definition.

Two themes that emerged from our analysis were people who (correctly) thought about privacy policies as descriptions of behaviors (e.g., "conditions under which your info will/will not be shared") versus people who misunderstood privacy policies as providing inherent protections (e.g., "the process they use to protect my personal data from others using it"). These two themes are consistent with prior work about how people interpret the term

"privacy policy" [74, 80–82]. However, our results also show distinct differences in the prevalence of these themes between different generations. Generation Z predominantly thought of privacy policies as descriptions of behaviors (12/15, compared to 3/15 responses that contained protection themes). By comparison, Generation R and Generation Alpha were more evenly split (5/11 vs. 4/11 and 3/14 vs. 5/14, respectively).

Two other themes emerged in how people talked about privacy policies, both with distinct generational divides. 4/14 Generation Alpha participants specifically talked about privacy policies in the context of the Internet ecosystem (e.g., "It is a thing that tells you warnings about the website"), language that did not appear in any of the definitions provided by Generation Z or Generation R. By contrast, 10/15 responses from Generation Z specifically associated privacy policies with companies or corporate behavior (e.g., "what the company can and cannot do with your data"). This type of framing occurred more rarely in Generation R (3/11) and not at all in Generation Alpha.

Cookie. Overall trends for the term "cookie" were similar to those observed for "privacy policy": it was relatively broadly understood, but there were definite distinctions in prevalence of correct understandings between different generations. 13/15 participants in Generation Z exhibited at least a vague understanding of the term "cookie" compared to 5/11 in Generation R and 2/11 in Generation Alpha. Misconceptions about cookies—including that they were ads or desserts—were particularly common among Generation Alpha.

The themes that emerged in our thematic analysis corresponded to different uses of cookies. The most common theme was data collection and tracking (e.g., "something that collects data about a user" or "electronic marker that allows company to track user's choices/path"). This theme was particularly prevalent among Generation Z (10/15 responses talked about something relating to this theme), but it also appeared in 2/14 Generation Alpha responses and 2/11 Generation R responses. Younger cohorts also talked thematically about advertising (2/14 Generation Alpha responses and 3/15 Generation Z responses), but this theme did not appear in any of the Generation R responses. Older cohorts sometimes talked about local storage or remembering state (4/15 in Generation Z and 1/11 in Generation R) or cookies acting as identifiers (2/15 in Generation Z and 1/11 in Generation R), but none of our Generation Alpha participants mentioned these themes. Additionally, a small number of participants in each cohort mentioned personalization (1/14 in Generation Alpha, 2/15 in Generation Z, and 1/11 in Generation R), and one Generation Alpha participant mentioned authentication.

Pixel Tag. Unlike the two previous terms, "pixel tag" was generally unknown to most participants across all three generational cohorts. Only three participants—all in Generation Z—had even a vague idea of what pixel tags were. 4/14 Generation Alpha participants held misconceptions about what a pixel tag was, which was rare in Generation Z (1/15) and Generation R (1/11). Most participants across all cohorts (10/14, 11/15, and 10/11 respectively) were aware that they did not know what the term "pixel tag" means. These results are consistent with prior work that has shown that the term "pixel tag" is not well understood [80], but the number of participants who did not know what the term meant was much higher than the 20% observed in that online study, likely reflecting lower levels of technical sophistication in our study population, particularly among children and older adults.

Given the small number of responses that said anything but "I don't know", our thematic analysis did not identify many themes. The one theme that did emerge was that the adult participants in Generation Z and Generation R who held misconceptions both conflated pixel tags with image pixels (e.g., "It is a point... as more pixels per square inch, the cleaner the picture. What it means for privacy and pixel tags escapes me"). This misconception was also identified in Tang et al.'s online study [80]. However, we found that none of the Generation Alpha participants with incorrect definitions held this misconception. Instead, the misconceptions exhibited by Generation Alpha were highly varied, including tabs ("Info in tabs.?.??.?") and redaction ("Hashtags covering any words that seem unfamiliar or inappropriate").

Cache. Most Generation Z participants (9/15) had at least a vague understanding that caches were a place for storing data. By contrast, this term was less broadly understood among Generation Alpha (2/14) and Generation R (2/11). Both Generation Alpha and Generation Z participants exhibited some misconceptions about the term, whereas most Generation R participants (9/11) simply stated that they did not know what the term meant.

Our thematic analysis found that many adult participants—10/15 in Generation Z and both Generation R participants who ventured a definition—defined caches in terms of data storage in some way. Four Generation Z participants provided more specific definitions that referenced temporary or local storage (e.g., "data that is stored in a device that is often used for background functions or often taking up space"), a theme that did not appear in responses from other generations. Conversely, Generation Alpha thought about caches in terms of functionality enabled. For example, 3/14 referenced offline access (e.g., "a thing that makes it so any website can run without an Internet connection").

Metadata. "Metadata", like all other terms, was understood best by Generation Z. 7/15 Generation Z participants had at least a vague idea of what the term means compared to 4/14 for Generation Alpha and 1/11 for Generation Z. None of our Generation Alpha participants fully understood the term or even understood the key idea behind the term. We also saw several misconceptions about this term across all generational cohorts (12/45). Generation R was most likely to recognize that they did not understand the term (7/11 compared to 6/14 in Generation Alpha and 3/15 in Generation Z). Our results for Generation Z were consistent with prior work conducted online, which found that 48% of that population could correctly define "metadata" [80]. Our results for Generation Alpha and Generation R show lower levels of understanding among those generational groups.

One theme that emerged was the use of concrete examples rather than more general definitions. 8/15 participants from Generation Z gave examples such as file size, image location, or the time when an image was taken, as did one member of Generation R. Associations with images were common (5/14 in Generation Alpha, 4/15 in Generation R, 1/11 in Generation Z), perhaps because participants had encountered the term in this context or perhaps because the privacy policy excerpt for this term mentioned a device's camera roll. Among incorrect responses, we observed two primary themes.

Three participants (one in Generation Alpha and two in Generation Z) associated the term "metadata" with the metasphere or the Internet more generally (e.g., "data that can be used to recreate physical environments virtually"). Two participants, both in Generation R, thought "metadata" referred to the amount of data (e.g., "large amount of data"). These themes were not identified by prior work, perhaps because they arose among generational cohorts that are not well-represented (or represented at all) in online studies.

Device Fingerprinting. We saw distinct generational differences in how participants understood the term "device fingerprinting". Among our Generation Z participants, 6/15 had at least a vague understanding of the term, a result that is consistent with prior online work [80]. However, only 1/14 in Generation Alpha and 1/11 in Generation R did. Miscomprehensions about the term were common among the younger cohorts (7/14 in Generation Alpha and 6/15 in Generation Z). By contrast, 10/11 participants in the retired cohort simply responded that they do not know what the term means.

Correct answers, by definition, all talked about the ability to identify devices or individuals in some way, for example, "a way for a website to identify your computer" or "How each device is different". This theme was most common among Generation Z (8/15), likely since this term was better understood by that generation. However it was also reflected in the single vaguely correct response from each other generation. The common theme that emerged among participants with miscomprehensions was conflation between device fingerprinting and fingerprint-based authentication. For example, "when you allow your fingerprint to be used to buy apps or unlock things". This theme, which was also observed in prior work [80], was most common among Generation Alpha (6/14), but two participants from Generation Z also reflected this theme. Other incorrect responses referred to digital forensics, stored passwords, or single sign-on.

Encryption. "Encryption" was one of the more commonly understood terms, with half of our participants demonstrating vague to complete comprehension. However, fewer participants fully understood the term. For example, a Generation Z participant defined encryption as: "Making something not available to the public; only available through password protection etc." We also observed generational differences, particularly between the two adult cohorts—in which a majority of participants provided a definition that was at least vaguely correct—and the children's cohort. Only 1/14 Generation Alpha children understood the big idea of encryption: they defined it as, "To scramble or make hard to understand and/or find".

Among correct responses, three general themes emerged from our thematic analysis: people thought of encryption in terms of encoding, protection, or scrambling. Most Generation Z participants thought of encryption either in terms of encoding (7/15, e.g., "a code that makes information illegible without a decryption device") or protection (7/15, e.g., "I think it means to hide so that other people can't see it"). For each encoding and protection, 2/11 Generation R responses also aligned with this theme. However, the most common theme among Generation R was scrambling. 3/11 Generation R responses exhibited this theme (e.g., "scrambling elements of message to make it unintelligible during transit"). One member of each of the Generation Z and Generation Alpha cohorts also used "scramble" or related language such as "jumble". There were no common themes among incorrect definitions, which varyingly thought encryption referred to data use, collected data, stored data, de-identified data, or identified data.

Biometric Data. Three primary themes emerged among definitions our participants provided for "biometric data". 5/15 participants in Generation Z referenced biometric identifiers such as facial ID or fingerprint-based authentication, e.g., "like face ID and finger prints and things related to your physical body". However, we did not observe this theme among responses from either of our other generations. A few people across each generation (1/14 in Generation Alpha, 3/15 in Generation Z, 2/11 in Generation R) described other physical attributes, e.g., "maybe data related to physical, such as height, weight, age etc." Four participants (2/14 in Generation Alpha, 1/15 in Generation Z, and 1/11 in Generation R) described personal information more broadly, e.g., "Data about you". Incorrect answers referred to survey data, search data, or device interactions.

Overall, we found that most people across all three generations struggled to correctly define technical terms. The most well understood term was "privacy policy", for which 9/15 Generation Z participants and 3/11 Generation R participants understood at least the big idea. For all other terms, no more than a third of our participants in any generation were able to correctly define the term. None of our participants understood the term "pixel tag". Overall, approximately 60% of responses to definition questions were "I don't know".

Generation Z generally understood terms better than either Generation Alpha or Generation R. Both Generation Alpha and Generation R demonstrated prevalent misunderstandings or lack of understanding for many of the privacy terms. None of our participants in Generation Alpha exhibited full understandings of any of the technical terms, and only two responses (one for cookies and one for encryption) demonstrated more than a vague understanding of any term. On average, Generation R responded that they did not know the meaning of a technical term 80% of the time.

4.3 Confidence with Privacy Terms

Subjects across age groups reported feeling confused or ignorant about the terms used in our study. A Generation R participant remarked that the survey "reaffirms for me the degree to which I feel used by these corporations". One of the Generation Z participants was so unfamiliar with some of the terms that they believed we had made them up ourselves. The majority of Generation Alpha also reported feeling confused, with several reporting that most of their answers were "I don't know". For 6 out of the 8 privacy terms studied, the majority of respondents expressed either being somewhat unconfident or very unconfident. No terms except "Privacy Policy" had a majority of Generation Alpha participants express confidence in their responses (8/14 of Generation Alpha responded as being "somewhat confident" in their privacy policy answers). Self-reported levels of confidence at defining various privacy terms are depicted in Figure 3.

The results from both our survey and our focus group showed that older adults feel overwhelmingly confused by privacy polices

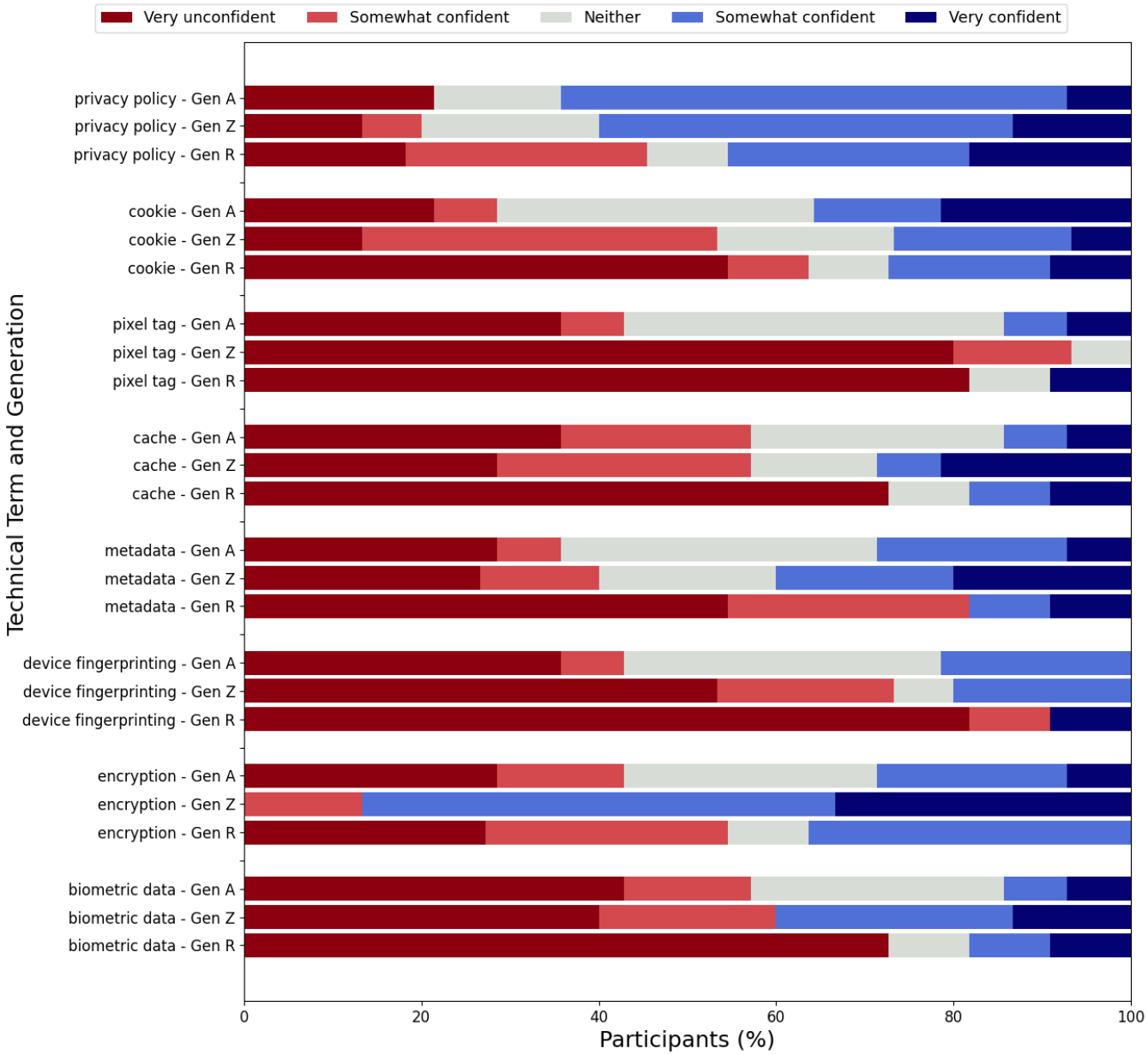


Figure 3: Self-reported confidence at defining various privacy terms across different generations.

and their terminology. They are aware of their lack of knowledge—Generation R reported much lower confidence than either Generation Z or Generation Alpha—but do not know who or where to turn to get answers. As one participant put it, "I don't even know where to go to find out what [these terms] mean". This is a stark contrast to Generation Alpha. After being prompted to discuss possible solutions to the lack of public explanation of what these terms meant, one Generation Alpha participant remarked that we should put up "[a] billboard on the highway that says 'just search it up'".

Not only are older adults aware of how little they understanding about these technical terms, they also feel bad about their lack of knowledge. Every Generation R participant agreed that taking the survey made them feel "stupid" and "really old". In addition, when they do search for answers to their online privacy and terminology questions, they are often brushed off or infantilized by those they ask—when defining what "cookie" meant, one participant said it was "something I was told to simply accept– no problem ([by] a computer technician)".

4.4 Implications and Directions for Improvement

When asked what the consequences of not knowing these privacy terms might be, responses fell into two high-level categories: potential consequences and current feelings.

Many participants across all generational cohorts talked about a fear of concrete consequences that could hypothetically arise from not knowing technical terms that appear in privacy policies. Two general themes emerged. Participants across all three generations mentioned fear of getting hacked as a possible consequence (3/15 in Generation Alpha, 4/15 in Generation Z, 4/11 in Generation R). For example, "Possible hackers or viruses, maybe even bad people finding info about you", "Probably getting hakt lol", and "maybe one can become subject to fraud or becoming a subject for hacking". Participants in the two adult generations (5/15 in Generation Z and 3/11 in Generation R) were also cognizant of possible privacy vulnerabilities. Many responses that exhibited this theme referred specifically to a lack of transparency and the consequent failure of informed consent. For example, a Generation Z participant said, "my data is probably being used and distributed much more than I may think and in ways that I do not know". Another Generation Z participant remarked, "the consequences could include unintentionally consenting to the usage/access of my personal data". A Generation R participant stated, "I think I am open to having personal information used for all manner of things that, if I actually knew, would be profoundly unsettling!" Other concrete hypothetical consequences included lack of access to features, loss of reputation, and the possibility of being sued.

Other participants talked about their current feelings as a consequence of not knowing privacy terms. Feelings of personal inadequacy were most salient for Generation R, whose responses included, "I feel very uninformed", "frustration [for] all who try to help me", and "the consequences are a feeling of being left behind on the shore of the 20th century". Generation Z shared feelings of dis-empowerment and frustration, for example, "Blindly giving your data without an understanding of these terms places power in the hands of the company. Without users having strong knowledge of these terms, companies have the ability to use said data with little oversight" and "even if I didn't agree with the privacy policy, I don't have the means to bargain or change the tenets of the policy". 6/14 Generation Alpha participants said they did not know what the consequences would be, but many acknowledged that there probably would be consequences. For example, one said, "Idk may be awful", and another said, "I think I should know".

When asked about privacy policy reform, participants across age groups were interested in a condensed version of privacy policies written in more common laymen's terms, with all 11 Generation R subjects, 7/15 Generation Z subjects, and 10/13[3] Generation Alpha subjects saying they would read privacy policies if they were condensed. However, one of the Generation Z participants remarked that "companies will not be incentivized to make them easier [to understand]".

[3]One of the 14 Generation Alpha participants excused themself to go to the bathroom and was not part of the focus group discussion.

5 DISCUSSION

Although this work was conducted with a non-representative, small-scale population, our results identify limitations of prior work and directions for future research. Based on these results, we make four concrete recommendations.

Recommendation 1: *Educators and non-profit organizations should develop and distribute generation-specific, targeted outreach and educational efforts for children and older adults.*

Our results show that while people across all generations struggle to understand technical terms, misunderstandings and lack of knowledge are particularly acute problems for children and older adults. People in these generational cohorts want to learn more about privacy terms, but for cohort specific reasons—Generation Alpha because it has not occurred to them, Generation R because they don't know where to look—currently-available educational materials are insufficient. Our results suggest that it would be valuable to develop generation-specific curricular modules and educational materials and to pursue targeted outreach efforts through elementary-level school visits and enrichment workshops at senior centers. One potential avenue could be gamified learning. Leech [62]—an adventure-style online game in which the player interacts with various other characters who discuss the privacy policy while journeying to the castle to take back their data—was found to be both enjoyable and educational. However, it was not designed for children, which might deter younger audiences, and it requires facility with computer games, which might deter older audiences. Generationally tailored—and publicly available—approaches could enhance data practice transparency for children and older adults.

Recommendation 2: *Researchers and software developers should design and evaluate generation-specific privacy-enhancing tools to address miscomprehensions relating to technical terms.*

People encounter technical terms in many different contexts, including privacy policies, consent interfaces, app privacy labels, and permission requests. In many of these contexts, comprehension of these terms is a precondition for informed consent, and yet research consistently finds that people do not understand the technical terms used in any of these contexts [25, 31, 45, 80, 97]. While some children may inherently be too young to develop nuanced understandings of complex technologies, users of all ages should be able to understand when described data use practices have potential privacy implications. To support this, usable privacy researchers and software developers should design and evaluate technical solutions for providing real-time information about misunderstood or unfamiliar terms. Possible approaches might include policy visualization tools (e.g., [30]), annotated policies, summaries, hyperlinks, or on-hover supplementary explanations and examples. These solutions need not be one-size-fits all. Just as content and advertising are personalized, privacy-enhancing tools could provide generation-specific guidance to enhance privacy without intruding unnecessarily on the user experience. However, all users should have access to all tools so that people whose needs deviate from the norm—or whose ages are incorrectly inferred—are not negatively impacted.

Recommendation 3: *Legislators and corporate privacy officers should pursue efforts to enhance privacy beyond transparency, notice and consent, and privacy self-management.*

While educational outreach and privacy-enhancing tools have potential to improve transparency by reducing misunderstandings and improving user comprehension of technical terms, those approaches are incremental improvements rather than full solutions. Our results extend a long line of prior work that shows that privacy disclosures are unreadable [2, 61], omit critical information [23, 85], and nudge people away from invoking their rights [18, 29, 60, 84, 87]. Moreover, notice and consent simply does not scale to the number of companies with which users regularly interact, nor does it handle the issue of users' inability to identifying the many third parties with access to their personal data [51, 70, 75, 76]. These results therefore provide further evidence that legislators should work to enhance privacy by introducing and enforcing legal requirements that move beyond enhancing transparency and focus instead on bring corporate data practices into alignmentment with cultural norms around data collection and use. Corporate privacy officers situated within companies should advocate for voluntary compliance with such social norms. Researchers can contributed by identifying and validating what current social norms are, by studying how such norms evolve over time, and by investigating the extent to which general social norms are consistent with the privacy needs of various subpopulations.

Recommendation 4: *The usable privacy research community should continue to invest in in-person, community-based research and avoid over-generalizing the results of online studies.*

Online crowd-sourcing platforms such as Amazon Mechanical Turk, CloudResearch, and Prolific are a convenient, fast, inexpensive way to recruit participants for both small-scale interview studies and large-scale survey and experimental user studies. It is therefore not surprising that an increasing number of user studies in the domain of usable privacy and security recruit their participants through these platforms [79]. However, these crowd-sourcing platforms are not fully representative of the overall population. Children under 18 are explicitly excluded from these platforms, and older adults are both underrepresented and poorly-represented—older adults registered on these platforms are significantly more technically sophisticated compared to the rest of their age cohort, resulting in online studies not being representative of older adults [68, 79]. Our results provide concrete evidence that online studies are missing generation-specific insights pertaining to children and older adults, and that the results of online studies do not consistently generalize to other age groups. While online studies are undeniably valuable, that work should be supplemented and replicated by in-person, community-based research that actively targets user populations who are excluded from online studies. Researchers should also seek to develop new recruitment methods—perhaps through collaborations with community organizations or with other researchers affiliated with less research-active institutions—that extend the scope of community-based research beyond our own personal communities.

6 CONCLUSION

Overall, we found that technical terms that appear in privacy policies are not understood or are misunderstood by many people across all three generational cohorts studied: children (Generation Alpha), young adults (Generation Z), and older adults (Generation R). However, we also observed differences between our three generational groups. Generation Z contained the highest proportion of participants who understood a term, while Generation Alpha and our constructed Generation R (consisting of both the Silent Generation and Baby Boomers) had lower proportions of participants who understood a concept or technical privacy term. We also identified varying themes in how our three generational cohorts understood technical terms.

Participants across all three generations reported that the survey and the associated terminology made them confused or frustrated, further illustrating the necessity of privacy policy reform to enable transparency about data use practices. However, Generation R participants more frequently said that they did not know what a term meant rather than venturing a guess, and Generation R reported the lowest level of confidence in defining these terms.

Our three generational cohorts differed in how they thought about the consequences of the lack of transparency about data practices that resulted from being unable to understand technical terms that appear in privacy policies. Generation R most commonly described the consequences of this knowledge gap in terms of feelings of personal inadequacy, and all Generation R participants agreed that their lack of knowledge made them feel stupid. Generation Z focused more on concrete security and privacy consequences, although several members of that generation described feelings of dis-empowerment. Generation Alpha was less cognizant of the implications of data collection and use.

To the best of our knowledge, this is the first study on privacy terms that has included Generation Alpha. Further research should continue to focus on this upcoming generation and their knowledge of the digital landscape they were born into. Much previous research with the Silent Generation or Baby Boomers has been conducted exclusively online, leading to a potentially biased sample. In addition to further research with Generation Alpha, more in-person research should be conducted with the Silent Generation and Baby Boomers to guarantee an accurate sample in distribution of technological comfort. Research should additionally be conducted for all generation groups in different geographic and socioeconomic areas to enable equitable privacy for people of all ages.

ACKNOWLEDGMENTS

This work was supported by in part by NSF grant 2317115 and by internal funds from Pomona College.

REFERENCES

[1] Andrick Adhikari, Sanchari Das, and Rinku Dewri. 2023. Evolution of Composition, Readability, and Structure of Privacy Policies over Two Decades. *Proceedings on Privacy Enhancing Technologies* 3 (2023), 138–153.

[2] Ryan Amos, Gunes Acar, Elena Lucherini, Mihir Kshirsagar, Arvind Narayanan, and Jonathan Mayer. 2021. Privacy Policies Over Time: Curation and Analysis of a Million-Document Dataset. In *Proceedings of the Web Conference*. 2165–2176.

[3] Annie Anton, Julia Earp, Qingfeng He, William Stufflebeam, Davide Bolchini, and Carlos Jensen. 2004. Financial Privacy Policies and the Need for Standardization. *IEEE Security & Privacy* 2 (2004), 36–45.

[4] Siddhant Arora, Henry Hosseini, Christine Utz, Vinayshekhar K. Bannihatti, Tristan Dhellemmes, Abhilasha Ravichander, Peter Story, Jasmine Mangat, Rex Chen, Martin Degeling, Tom Norton, Thomas Hupperich, Shomir Wilson, and Norman Sadeh. 2022. A Tale of Two Regulatory Regimes: Creation and Analysis of a Bilingual Privacy Policy Corpus. In *Proceedings of the 13th Conference on Language Resources and Evaluation*. 5460–5472.

[5] Article 29 Working Party. 2017. Guidelines on Transparency under Regulation 2016/679.

[6] Brooke Auxier, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner. 2019. Americans' Attitudes and Experiences with Privacy Policies and Laws. *Pew Research Center: Internet, Science & Tech* (2019).

[7] Solon Barocas and Helen Nissenbaum. 2009. On Notice: The Trouble with Notice and Consent. In *Proceedings of the Engaging Data Forum: The First International Forum on the Application and Management of Personal Electronic Information*.

[8] Peter Breese and William Burman. 2005. Readability of Notice of Privacy Forms Used by Major Healthcare Institutions. *The Journal of the American Medical Association* 293 (2005), 1588–1594.

[9] Rochelle Cadogan. 2011. An Imbalance Of Power: The Readability Of Internet Privacy Policies. *Journal of Business & Economics Research* 2 (2011), 48–61.

[10] Rex Chen, Fei Fang, Thomas Norton, Aleecia M. McDonald, and Norman Sadeh. 2021. Fighting the Fog: Evaluating the Clarity of Privacy Disclosures in the Age of CCPA. In *Proceedings of the 20th Workshop on Workshop on Privacy in the Electronic Society*. 73–102.

[11] Committee for Protection of Human Subjects, University of California Berkeley. 2022. Compensation of Research Subjects. https://cphs.berkeley.edu/compensation.pdf. Accessed on February 29, 2024.

[12] Computer Security Resource Center. 2024. Glossary. https://csrc.nist.gov/glossary. Accessed on October 4, 2022.

[13] Gitanjali Das, Cynthia Cheung, Camille Nebeker, Matthew Bietz, and Cinnamon Bloss. 2018. Privacy Policies for Apps Targeted Toward Youth: Descriptive Analysis of Readability. *JMIR Mhealth Uhealth* 6, 1 (2018).

[14] Martin Degeling, Christine Utz, Christopher Lentzsch, Henry Hosseini, Florian Schaub, and Thorsten Holz. 2019. We Value Your Privacy... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy. In *Proceedings of the Network and Distributed System Security Symposium*. 1–15.

[15] John Dempsey, Gavin Sim, Brendan Cassidy, and Vihn-Thong Ta. 2021. Children Designing Privacy Warnings: Informing a Set of Design Guidelines. *International Journal of Child-Computer Interaction* (2021).

[16] Ratan Dey, Yuan Ding, and Keith W. Ross. 2013. Profiling High-school Students with Facebook: How Online Privacy Laws Can Actually Increase Minors' Risk. In *Proceedings of the 2013 Conference on Internet Measurement Conference*. 405–416.

[17] Michael Dimock. 2019. Defining Generations: Where Millennials End and Generation Z Begins. *Pew Research Center* 17, 1 (2019), 1–7.

[18] Nora A. Draper and Joseph Turow. 2019. The Corporate Cultivation of Digital Resignation. *New Media & Society* 21, 8 (2019), 1824–1839.

[19] Isioma Elueze and Anabel Quan-Haase. 2018. Privacy Attitudes and Concerns in the Digital Lives of Older Adults: Westin's Privacy Attitude Typology Revisited. *American Behavioral Scientist* 62, 10 (2018), 1372–1391.

[20] Tatiana Ermakova, Benjamin Fabian, and Eleonora Babina. 2015. Readability of Privacy Policies of Healthcare Websites. In *12th International Conference on Wirtschaftsinformatik*. 1085–1099.

[21] The European Parliament and the Council of the European Union. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Wersons with regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), L 119/1.

[22] Benjamin Fabian, Tatiana Ermakova, and Tino Lentz. 2017. Large-Scale Readability Analysis of Privacy Policies. In *Proceedings of the International Conference on Web Intelligence*. 18–25.

[23] Ming Fan, Le Yu, Sen Chen, Hao Zhou, Xiapu Luo, Shuyue Li, Yang Liu, Jun Liu, and Ting Liu. 2020. An Empirical Evaluation of GDPR Compliance Violations in Android mHealth Apps. In *31st International Symposium on Software Reliability Engineering*. 253–264.

[24] Michelle Faverio. 2022. Share of Those 65 and Older Who Are Tech Users Has Grown in the Past Decade. https://www.pewresearch.org/fact-tank/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/. Accessed on October 30, 2023.

[25] Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. 2012. Android Permissions: User Attention, Comprehension, and Behavior. In *Proceedings of the 8th Symposium on Usable Privacy and Security*.

[26] Rudolph Flesch. 1948. A New Readability Yardstick. *Journal of Applied Psychology* 32, 3 (1948), 221–233.

[27] Alisa Frik, Leysan Nurgalieva, Joyce Lee, Florian Schaub, and Serge Egelman. 2019. Privacy and Security Threat Models and Mitigation Strategies of Older Adults. In *15th Symposium on Usable Privacy and Security*. 21–40.

[28] Mark A. Graber, Donna M. D'Alessandro, and Jill Johnson-West. 2002. Reading Level of Privacy Policies on Internet Health Web Sites. *Journal of Family Practice* 51, 7 (2002), 642–642.

[29] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L Toombs. 2018. The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.

[30] Wentao Guo, Jay Rodolitz, and Eleanor Birrell. 2020. Poli-see: An Interactive Tool for Visualizing Privacy Policies. In *Proceedings of the 19th Workshop on Privacy in the Electronic Society*. 57–71.

[31] Hana Habib, Megan Li, Ellie Young, and Lorrie Faith Cranor. 2022. "Okay, whatever": An Evaluation of Cookie Consent Interfaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–27.

[32] Hana Habib, Sarah Pearman, Jiamin Wang, Yixin Zou, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh, and Florian Schaub. 2020. "It's a scavenger hunt": Usability of Websites' Opt-Out and Data Deletion Choices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[33] Alexis Hiniker, Amelia Wang, Jonathan Tran, Mingrui Ray Zhang, Jenny Radesky, Kiley Sobel, and Sungsoo Ray Hong. 2021. Can Conversational Agents Change the Way Children Talk to People?. In *Proceedings of the 20th Annual ACM Interaction Design and Children Conference*. 338–349.

[34] Mark Hochhauser. 2001. Lost in the Fine Print: Readability of Financial Privacy Notices. https://privacyrights.org/resources/lost-fine-print-readability-financial-privacy-notices-hochhauser. Accessed on October 30, 2023.

[35] Musa Jafar and Amjad Abdullat. 2011. Exploratory Analysis Of The Readability Of Information Privacy Statement Of The Primary Social Networks. *Journal of Business & Economics Research* 7, 12 (2011), 123–142.

[36] Mengtian Jiang, Hsin-yi Sandy Tsai, Shelia R. Cotten, Nora J. Rifon, Robert LaRose, and Saleem Alhabash. 2016. Generational Differences in Online Safety Perceptions, Knowledge, and Practices. *Educational Gerontology* 42, 9 (2016), 621–634.

[37] Lisa M. Jones, Kimberly J. Mitchell, and Cheryl L. Beseler. 2023. The Impact of Youth Digital Citizenship Education: Insights from a Cluster Randomized Controlled Trial Outcome Evaluation of the Be Internet Awesome (BIA) Curriculum. *Journal of Contemporary School Psychology* (2023).

[38] John Koetsier. 2023. Apple Sold 8 Of The Top 10 Best-Selling Phones Of 2022. https://www.forbes.com/sites/johnkoetsier/2023/03/07/apple-sold-8-of-the-top-10-best-selling-phones-of-2022/ Accessed on October 30, 2023.

[39] Barbara Krumay and Jennifer Klar. 2020. Readability of Privacy Policies. In *IFIP Annual Conference on Data and Applications Security and Privacy*. 388–399.

[40] Priya Kumar, Shalmali Milind Naik, Utkarsha Ramesh Devkar, Marshini Chetty, Tamara L. Clegg, and Jessica Vitak. 2017. 'No Telling Passcodes Out Because They're Private': Understanding Children's Mental Models of Privacy and Security Online. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–21.

[41] Renuka Kumar, Apurva Virkud, Ram Sundara Raman, Atul Prakash, and Roya Ensafi. 2022. A Large-Scale Investigation into Geodifferences in Mobile Apps. In *31st USENIX Security Symposium*. 1203–1220.

[42] Dominika Kuźnicka-Błaszkowska. 2022. Protecting Children's Personal Data under General Data Protection Regulation and California Consumer Privacy Act in Relation to Information Society Services–European Perspective. *Przegląd Prawa Konstytucyjnego* 70, 6 (2022), 487–498.

[43] Alexis R. Lauricella, Drew P. Cingel, Leanne Beaudoin-Ryan, Michael B. Robb, Melissa Saphir, and Ellen A. Wartella. 2016. The Common Sense Census: Plugged-in Parents of Tweens and Teens. *San Francisco, CA: Common Sense Media* (2016).

[44] Stephen D. Lewis, Robert G. Colvard, and C. N. Adams. 2008. A Comparison of the Readability of Privacy Statements of Banks, Credit Counseling Companies, and Check Cashing Companies. *Journal of Organizational Culture, Communications and Conflict* 12, 2 (2008), 87–93.

[45] Yanzi Lin, Jaideep Juneja, Eleanor Birrell, and Lorrie Faith Cranor. 2024. Data Safety vs. App Privacy: Comparing the Usability of Android and iOS Privacy Labels. *Proceedings on Privacy Enhancing Technologies* 2 (2024).

[46] Thomas Linden, Rishabh Khandelwal, Hamza Harkous, and Kassem Fawaz. 2020. The Privacy Policy Landscape after the GDPR. *Proceedings on Privacy Enhancing Technologies* 2020, 1 (2020), 47–64.

[47] Sonia Livingstone and Magdalena Bober. 2004. UK Children Go Online: Surveying the Experiences of Young People and their Parents. http://eprints.lse.ac.uk/395/1/UKCGOsurveyreport.pdf.

[48] Juniper Lovato, Philip Mueller, Parisa Suchdev, and Peter Dodds. 2023. More Data Types More Problems: A Temporal Analysis of Complexity, Stability, and Sensitivity in Privacy Policies. In *Proceedings of the 2023 ACM Conference on*

Fairness, Accountability, and Transparency. 1088–1100.
[49] Sunil Manandhar, Kaushal Kafle, Benjamin Andow, Kapil Singh, and Adwait Nadkarni. 2022. Smart Home Privacy Policies Demystified: A Study of Availability, Content, and Coverage. In *31st USENIX Security Symposium*. 3521–3538.
[50] Mark McCrindle and Ashley Fell. 2020. Understanding Generation Alpha. https://generationalpha.com/wp-content/uploads/2020/02/Understanding-Generation-Alpha-McCrindle.pdf. Accessed on October 30, 2023.
[51] Aleecia M. McDonald and Lorrie Faith Cranor. 2008. The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Policy for the Information Society* 4, 3 (2008), 543–568.
[52] Gabriele Meiselwitz. 2013. Readability Assessment of Policies and Procedures of Social Networking Sites. In *Online Communities and Social Computing*. 67–75.
[53] George R. Milne, Mary J. Culnan, and Henry Greene. 2006. A Longitudinal Assessment of Online Privacy Notice Readability. *Journal of Public Policy & Marketing* 25, 2 (2006), 238–249.
[54] Caroline Lancelot Miltgen and Dominique Peyrat-Guillard. 2014. Cultural and Generational Influences on Privacy Concerns: A Qualitative Study in Seven European Countries. *European Journal of Information Systems* 23, 2 (2014), 103–125.
[55] Victor Morel and Raúl Pardo. 2020. SoK: Three Facets of Privacy Policies. In *Proceedings of the 19th Workshop on Privacy in the Electronic Society*. 41–56.
[56] Savanthi Murthy, Karthik S. Bhat, Sauvik Das, and Neha Kumar. 2021. Individually Vulnerable, Collectively Safe: The Security and Privacy Practices of Households with Older Adults. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–24.
[57] Jason M. Nagata, Catherine A. Cortez, Chloe J. Cattle, Kyle T. Ganson, Puja Iyer, Kirsten Bibbins-Domingo, and Fiona C. Baker. 2022. Screen Time Use Among U.S. Adolescents During the COVID-19 Pandemic: Findings from the Adolescent Brain Cognitive Development (ABCD) Study. *JAMA pediatrics* 176, 1 (2022), 94–96.
[58] National Center for Education Statistics. 2023. *Children's Internet Access at Home*. https://nces.ed.gov/programs/coe/indicator/cch/home-internet-access Accessed on December 8, 2023.
[59] Jonathan A. Obar and Anne Oeldorf-Hirsch. 2020. The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services. *Information, Communication & Society* 23, 1 (2020), 128–147.
[60] Sean O'Connor, Ryan Nurwono, Aden Siebel, and Eleanor Birrell. 2021. (Un)clear and (In)conspicuous: The Right to Opt-Out of Sale Under CCPA. In *Workshop on Privacy in the Electronic Society*. 59–72.
[61] Junhyoung Oh, Jinhyoung Hong, Changsoo Lee, Jemin Justin Lee, Simon S. Woo, and Kyungho Lee. 2021. Will EU's GDPR Act as an Effective Enforcer to Gain Consent? *IEEE Access* 9 (2021), 79477–79490.
[62] Sebastian Pape, Alexander Klauer, and Michaela Rebler. 2021. Leech: Let's Expose Evidently bad data Collecting Habits—Towards a Serious Game on Understanding Privacy Policies (Poster). USENIX Symposium on Usable Privacy and Security.
[63] PayScale. 2022. Cost of Living in Claremont, California. https://www.payscale.com/cost-of-living-calculator/California-Claremont. Accessed on December 10, 2022.
[64] Irene Pollach. 2005. A Typology of Communicative Strategies in Online Privacy Policies: Ethics, Power and Informed Consent. *Journal of Business Ethics* 62 (12 2005), 221–235.
[65] Irene Pollach. 2007. What's Wrong with Online Privacy Policies? *Commun. ACM* 50, 9 (2007), 103–108.
[66] Robert W. Proctor, M. Athar Ali, and Kim-Phuong L. Vu. 2008. Examining Usability of Web Privacy Policies. *International Journal of Human-Computer Interaction* 24, 3 (2008), 307–328.
[67] Hirak Ray, Flynn Wolf, Ravi Kuber, and Adam J. Aviv. 2019. "Woe is me:" Examining Older Adults' Perceptions of Privacy. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–6.
[68] Elissa M. Redmiles, Sean Kross, and Michelle L. Mazurek. 2019. How Well Do My Results Generalize? Comparing Security and Privacy Survey Results from mTurk, Web, and Telephone Samples. In *IEEE Symposium on Security and Privacy*. 1326–1343.
[69] Joel R. Reidenberg, Travis Breaux, Lorrie Faith Cranor, Brian French, Amanda Grannis, James T. Graves, Fei Liu, Aleecia McDonald, Thomas B. Norton, and Rohan Ramanath. 2015. Disagreeable Privacy Policies: Mismatches between Meaning and Users' Understanding. *Berkeley Technology Law Journal* 30 (2015), 39–68.
[70] Neil Richards and Woodrow Hartzog. 2018. The Pathologies of Digital Consent. *Washington University Law Review* 96 (2018), 1461–1504.
[71] Victoria Rideout and Michael B. Robb. 2017. The Common Sense Census: Media Use by Kids Age Zero to Eight. *San Francisco, CA: Common Sense Media* 263 (2017).
[72] Julie Robillard, Tanya L. Feng, Arlo B. Sporn, Jen-Ai Lai, Cody Lo, Monica Ta, and Roland Nadler. 2019. Availability, Readability, and Content of Privacy Policies and Terms of Agreements of Mental Health Apps. *Internet Interventions* 17 (2019).
[73] Ravi Inder Singh, Manasa Sumeeth, and James Miller. 2011. A User-centric Evaluation of the Feadability of Privacy Policies in Popular Web Sites. *Information Systems Frontiers* 13, 4 (2011), 501–514.
[74] Aaron Smith. 2014. What Internet Users Know about Technology and the Web. https://policycommons.net/artifacts/619447/what-internet-users-know-about-technology-and-the-web/1600546/. Accessed on December 10, 2022.
[75] Daniel J. Solove. 2013. Introduction: Privacy Self-Management and the Consent Dilemma. *Harvard Law Review* 126 (2013), 1880–1903.
[76] Daniel J. Solove. 2021. The Myth of the Privacy Paradox. *George Washington Law Review* 89 (2021), 1–51.
[77] StatCounter. 2023. Search Engine Market Share United States Of America Sept 2022 - Sept 2023. https://gs.statcounter.com/search-engine-market-share/all/united-states-of-america Accessed on October 30, 2023.
[78] Statista. 2023. Social Media Usage in the United States. https://www.statista.com/study/40227/social-social-media-usage-in-the-united-states-statista-dossier/ Accessed on October 30, 2023.
[79] Jenny Tang, Eleanor Birrell, and Ada Lerner. 2022. Replication: How Well Do My Results Generalize Now? The External Validity of Online Privacy and Security Surveys. In *18th Symposium on Usable Privacy and Security*. 367–385.
[80] Jenny Tang, Hannah Shoemaker, Ada Lerner, and Eleanor Birrell. 2021. Defining Privacy: How Users Interpret Technical Terms in Privacy Policies. *Proceedings on Privacy Enhancing Technologies* 2021, 3 (2021), 70–94.
[81] Joseph Turow, Lauren Feldman, and Kimberly Meltzer. 2005. Open to Exploitation: America's Shoppers Online and Offline. A Report from the Annenberg Public Policy Center of the University of Pennsylvania.
[82] Joseph Turow, Michael Hennessy, and Nora Draper. 2018. Persistent Misperceptions: Americans' Misplaced Confidence in Privacy Policies, 2003–2015. *Journal of Broadcasting & Electronic Media* 62, 3 (2018), 461–478.
[83] United States Code. 1998. Children's Online Privacy Protection Act of 1998, 15 U.S.C. 6501–6505.
[84] Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. 2019. (Un)informed Consent: Studying GDPR Consent Notices in the Field. In *ACM SIGSAC Conference on Computer and Communications Security*. 973–990.
[85] Evangelia Vanezi, George Zampa, Christos Mettouris, Alexandros Yeratziotis, and George A Papadopoulos. 2021. CompLicy: Evaluating the GDPR Alignment of Privacy Policies—A Study on Web Platforms. In *International Conference on Research Challenges in Information Science*. 152–168.
[86] Kim-Phuong L. Vu, Vanessa Chambers, Fredrick P. Garcia, Beth Creekmur, John Sulaitis, Deborah Nelson, Russell Pierce, and Robert W. Proctor. 2007. How Users Read and Comprehend Privacy Policies. In *Human Interface and the Management of Information. Interacting in Information Environments*. 802–811.
[87] Ari Ezra Waldman. 2020. Cognitive Biases, Dark Patterns, and the 'Privacy Paradox'. *Current Opinion in Psychology* 31 (2020), 105–109.
[88] Charles Weir, Ben Hermann, and Sascha Fahl. 2020. From Needs to Actions to Secure Apps? The Effect of Requirements and Developer Practices on App Security. In *29th USENIX Security Symposium*. 289–305.
[89] David Wendler, Jonathan E. Rackoff, Ezekiel J. Emanuel, and Christine Grady. 2002. The Ethics of Paying for Children's Participation in Research. *The Journal of Pediatrics* 141, 2 (2002), 166–171.
[90] Alan F. Westin. 1967. *Privacy and Freedom*. Athenum, New York.
[91] Wikipedia. 2022. Cache (computing). https://en.wikipedia.org/wiki/Cache_(computing). Accessed on October 4, 2022.
[92] Wikipedia. 2022. Device Fingerprint. https://en.wikipedia.org/wiki/Device_fingerprint. Accessed on October 4, 2022.
[93] Wikipedia. 2022. Privacy Policy. https://en.wikipedia.org/wiki/Privacy_policy. Accessed on October 4, 2022.
[94] Wikipedia. 2022. Web Beacon. https://en.wikipedia.org/wiki/Web_beacon#cite_note-1. Accessed on October 4, 2022.
[95] Jason C. Yip, Kiley Sobel, Xin Gao, Allison Marie Hishikawa, Alexis Lim, Laura Meng, Romaine Flor Ofiana, Justin Park, and Alexis Hiniker. 2019. Laughing is Scary, but Farting is Cute: A Conceptual Model of Children's Perspectives of Creepy Technologies. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–15.
[96] Razieh Nokhbeh Zaeem and K. Suzanne Barber. 2020. The Effect of the GDPR on Privacy Policies: Recent Progress and Future Promise. *ACM Transactions on Management Information Systems* 12, 1 (2020), 1–20.
[97] Shikun Zhang, Yuanyuan Feng, Yaxing Yao, Lorrie Faith Cranor, and Norman Sadeh. 2022. How Usable Are iOS App Privacy Labels? *Proceedings on Privacy Enhancing Technologies* 4 (2022), 204–228.
[98] Leah Zhang-Kennedy, Christine Mekhail, Yomna Abdelaziz, and Sonia Chiasson. 2016. From Nosy Little Brothers to Stranger-Danger: Children and Parents' Perception of Mobile Threats. In *Proceedings of the The 15th International Conference on Interaction Design and Children*. 388–399.
[99] Jun Zhao, Ge Wang, Carys Dally, Petr Slovak, Julian Edbrooke-Childs, Max Van Kleek, and Nigel Shadbolt. 2019. 'I Make Up a Silly Name': Understanding Children's Perception of Privacy Risks Online. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

A STUDY PROTOCOL

(1) What year where you born? [free response]
(2) What is your gender?
  - Girl/Woman
  - Boy/Man
  - Non-binary
  - Other/I prefer to self-describe
(3) What is your racial identity (check as many that apply)?
  - White (European)
  - Black
  - East Asian
  - South Asian
  - Middle Eastern/West Asian
  - Native American
  - Native Pacific Islander
  - Latino

Imagine you are signing up for a new website or app. When you create your username and password, you are asked to agree to the privacy policy, and as you read through it, you notice it uses some technological words that are not used in everyday conversation. Please answer the following questions on what these words mean to the best of your ability.

An excerpt from a Google, Facebook, or Apple privacy policy is provided to help give context to each term. If you do not know what it means in this context, you can answer "I don't know".

(4) Excerpt: "In addition to this Privacy Policy, we provide data and privacy information embedded in our products and certain features that ask to use your personal information. This product-specific information is accompanied by our Data & Privacy Icon." What is a "privacy policy"? [free response]
(5) How confident are you that you provided the correct definition of "privacy policy"?
  - Very unconfident
  - Somewhat unconfident
  - Neither confident nor unconfident
  - Somewhat confident
  - Very confident
(6) Excerpt: "If you prefer that Apple not use cookies, we provide you with the means to disable their use. If you want to disable cookies and you're using the Safari web browser, choose "Block all cookies" in Safari's privacy settings." What is a "cookie"? [free response]
(7) How confident are you that you provided the correct definition of "cookie"?
  - Very unconfident
  - Somewhat unconfident
  - Neither confident nor unconfident
  - Somewhat confident
  - Very confident
(8) Excerpt: "Pixel tags are often used in combination with cookies." What are "pixel tags"? [free response]
(9) How confident are you that you provided the correct definition of "pixel tags"?
  - Very unconfident
  - Somewhat unconfident
  - Neither confident nor unconfident
  - Somewhat confident
  - Very confident
(10) Excerpt: "[a cache] can, for example, enable a web application to run without an internet connection." What is a "cache"? [free response]
(11) How confident are you that you provided the correct definition of "cache"?
  - Very unconfident
  - Somewhat unconfident
  - Neither confident nor unconfident
  - Somewhat confident
  - Very confident
(12) Excerpt: "For example, if you give us permission to access your device's camera roll, we collect metadata." What is "metadata"? [free response]
(13) How confident are you that you provided the correct definition of "metadata"?
  - Very unconfident
  - Somewhat unconfident
  - Neither confident nor unconfident
  - Somewhat confident
  - Very confident
(14) Excerpt: "The following policy applies to all sites and/or applications using Google Analytics and/or Google Analytics for Firebase: You must not use device fingerprints or locally shared objects (e.g. Flash cookies, Browser Helper Objects, HTML5 local storage) other than HTTP cookies." What is "device fingerprinting"? [free response]
(15) How confident are you that you provided the correct definition of "device fingerprinting"?
  - Very unconfident
  - Somewhat unconfident
  - Neither confident nor unconfident
  - Somewhat confident
  - Very confident
(16) Excerpt: "For example, we encrypt your information when it's in transit over public networks." What is "encryption"? [free response]
(17) How confident are you that you provided the correct definition of "encryption"?
  - Very unconfident
  - Somewhat unconfident
  - Neither confident nor unconfident
  - Somewhat confident
  - Very confident
(18) Excerpt: "Categories of personal information we collect: Biometric data, if you choose to provide it, such as [x] in Google's product development studies." What is "biometric data"? [free response]
(19) How confident are you that you provided the correct definition of "biometric data"?
  - Very unconfident
  - Somewhat unconfident
  - Neither confident nor unconfident
  - Somewhat confident
  - Very confident
(20) Overall, how confident do you feel with your responses?

- I think I got most/all of my answers correct
- I think I got some answers right and some wrong
- I think I got most/all of my answers wrong
- some of my answers were "I don't know", but I think I got most/all of the other answers correct
- some of my answers were "I don't know", but I think I got some of the other answers correct and some wrong
- some of my answers were "I don't know", but I think I got most/all of the other answers wrong
- most of my answers were "I don't know"

(21) How comfortable are you with computers?
- Very uncomfortable
- Somewhat uncomfortable
- Neither comfortable nor uncomfortable
- Somewhat comfortable
- Very comfortable

(22) How comfortable are you with smartphones/tablets?
- Very uncomfortable
- Somewhat uncomfortable
- Neither comfortable nor uncomfortable
- Somewhat comfortable
- Very comfortable

(23) After taking this survey, how interested are you in learning what these terms actually mean?
- Very uninterested
- Somewhat uninterested
- Neither interested nor uninterested
- Somewhat interested
- Very interested

(24) Lastly, if any, what do you think the consequences are of not knowing what these terms mean? [free response]

B CORRECT DEFINITIONS

Privacy Policy: "A statement or legal document (in privacy law) that discloses some or all of the ways a party gathers, uses, discloses, and manages a customer or client's data." [93]

Cookie: "A piece of state information supplied by a Web server to a browser, in a response for a requested resource, for the browser to store temporarily and return to the server on any subsequent visits or requests." [12]

Pixel Tags: "A technique used on web pages and email to unobtrusively (usually invisibly) allow checking that a user has accessed some content. Web beacons are typically used by third parties to monitor the activity of users at a website for the purpose of web analytics or page tagging." [94]

Cache: "A hardware or software component that stores data so that future requests for that data can be served faster; the data stored in a cache might be the result of an earlier computation or a copy of data stored elsewhere." [91]

Metadata: "Information describing the characteristics of data including, for example, structural metadata describing data structures (e.g., data format, syntax, and semantics) and descriptive metadata describing data contents (e.g., information security labels)." [12]

Device Fingerprinting: "Information collected about the software and hardware of a remote computing device for the purpose of identification... Device fingerprints can be used to fully or partially identify individual devices even when persistent cookies (and zombie cookies) cannot be read or stored in the browser, the client IP address is hidden, or one switches to another browser on the same device." [92]

Encryption: "Cryptographic transformation of data (called "plaintext") into a form (called "ciphertext") that conceals the data's original meaning to prevent it from being known or used. If the transformation is reversible, the corresponding reversal process is called "decryption," which is a transformation that restores encrypted data to its original state." [12]

Biometric Data: "Biological attribute of an individual from which distinctive and repeatable values can be extracted for the purpose of automated recognition. Fingerprint ridge structure and face topography are examples of biometric characteristics." [12]