Contextualizing Interpersonal Data Sharing in Smart Homes

Weijia He
Dartmouth College
weijia.he@dartmouth.edu

Nathan Reitinger
University of Maryland, College Park
nlr@umd.edu

Atheer Almogbil
Johns Hopkins University
aalmogb1@jhu.edu

Yi-Shyuan Chiang
University of Illinois
Urbana-Champaign
ysc6@illinois.edu

Timothy J. Pierson
Dartmouth College
timothy.j.pierson@dartmouth.edu

David Kotz
Dartmouth College
david.f.kotz@dartmouth.edu

ABSTRACT

A key feature of smart home devices is monitoring the environment and recording data. These devices provide security via motion-detection video alerts, cost-savings via thermostat usage history, and peace of mind via functions like auto-locking doors or water leak detectors. At the same time, the sharing of this information in interpersonal relationships—though necessary—is currently accomplished on an all-or-nothing basis. This can easily lead to oversharing in a multi-user environment. Although prior work has studied people's perceptions of information sharing with vendors or ISPs, the sharing of household data among users who interact personally is less well understood. Interpersonal situations make data sharing much more context-based and, thus, more complicated. In this paper, we use themes from the theory of contextual integrity in an online survey ($n = 1,992$) to study how people perceive data sharing with others in smart homes and inform future designs and research. Our results show that data recipients in a smart home can be reduced to three major groups, and data types matter more than device types. We also found that the types of access control desired by users can vary from scenario to scenario. Depending on whom they are sharing data with and about what data, participants expressed varying levels of comfort when presented with different types of access control (e.g., explicit approval versus time-limited access). Taken together, this provides strong evidence that a more dynamic access control system is needed, and we can design it in a more usable way.

KEYWORDS

IoT, Smart home, Privacy, Access control, Contextual Integrity

1 INTRODUCTION

Individuals can use web portals or smartphone apps to view the current status and recent events captured by smart devices in their home, including the current indoor temperature, whether lights are on or off, captured video footage, and past conversations with a voice assistant. Monitoring their home enables them to understand their daily routines better, confirm everything is in order while away, and maintain accountability if something goes wrong [28].

When multiple people share a smart home—whether as residents, landlords, or guests—equal access to the data may raise privacy concerns. There have been several reports about how smart home devices are turned into surveillance devices, enabling abuse and stalking [46]. Blocking everyone's access to smart-home data is also not practical [47]. For example, when smart home devices are shared between romantic partners or roommates, one may expect all device data to be equally accessible. Similarly, babysitters, a handyperson, and in-home caregivers may need the data to perform their tasks more effectively. Likewise, in the event of robbery and theft, it may be helpful or necessary to share video footage with law enforcement or insurance companies as evidence.

This potentially leaves us in an uncomfortable scenario: either share all data with someone (complete privacy loss) or share no data with them (complete utility loss). Many studies explore users' perceptions of institutional entities that receive smart home data, such as manufacturers, ISPs, and governments [7, 9, 38, 52]. These parties, although often essential, never have a personal connection with the smart home users. Users' perceptions and expectations of them may differ from those of other smart home users, with whom the owner may personally interact. Others that studied data sharing with individuals personally known to the resident, on the other hand, often focus narrowly on a particular device (e.g., voice assistants) or population (e.g., Airbnb, domestic workers, elders, and so forth) [3, 34]. They fail to put these scenarios in a more general setting, where multiple relationships, devices, and contexts co-exist. With big tech companies building cross-vendor platforms for smart homes (e.g., Google Home [21], Samsung's SmartThings [42], Apple's HomeKit [6]), a smart home system needs to accommodate and shift settings between different scenarios, which requires a more holistic privacy setting or access control mechanism.

Understanding varying contexts of users' preferences and decisions is the first step towards building a holistic usable system, but it is challenging, as a context often contains multiple *factors* that are at play during one's decision process. Prior work has demonstrated the need for more expressive context-aware access control systems that can encompass different scenarios in a smart home [13, 23, 38]. Unfortunately, these systems often face usability challenges in real life, where people find them too complicated to use [51].

Therefore, this paper aims to inform a more efficient and effective design of data-sharing mechanisms for smart homes by exploring people's preferences for data sharing in smart homes and how their preferences change based on varying contexts (e.g., data recipients, data types, and sharing principles). To analyze these questions, we designed a vignette survey

Proceedings on Privacy Enhancing Technologies 2024(2), 295–312

https://doi.org/10.56553/popets-2024-0051

based on the Contextual Integrity (CI) privacy framework [39]. Specifically, we used the survey to answer three main research questions:

RQ1: How do smart-home users perceive the sharing of smart-home data with different people?
RQ2: Do smart-home users' preferences for data sharing depend more on device types or data types? How would device and data types change users' opinions?
RQ3: What access control mechanism can support users' preferences about data sharing in various situations?

We use the CI framework to disambiguate these contexts. We chose the CI framework because it is a well-established framework that breaks a context down into a set of parameters, making it easier to analyze the situation systematically [39]. Using the theory of CI, we identify and decouple three different factors that are relevant to data sharing in a smart home: (1) data recipient (i.e., who will receive the data), (2) data type (i.e., what type of data is shared), and (3) sharing principles (i.e., under what circumstances would the data be shared). For each factor, we made a list of potential variants, relying on past literature, actual home IoT devices, and market statistics. This led us to 2,178 combinations, including 11 types of data recipients, 18 types of data from 6 different smart home devices, and 11 types of access control mechanisms. We recruited 1,992 participants from Prolific [26].

We found that data recipients are the most crucial factor in our participants' decision process. Based on participants' responses, the 11 types of data recipients can be roughly divided into three categories: *long-term residents* (spouses/significant others, kids, and roommates), *domestic workers* (in-home caretaker, handyperson, babysitter), and *incidental users* (guests, local law enforcement, landlords, neighbors, and insurance representatives). Participants are generally willing to share data with all the long-term residents. They may be hesitant to share data with domestic workers, but they are still more comfortable sharing data with them than with incidental users. In other words, it seems like participants are willing to share more with longer-term relationships rather than transient ones.

On the device and data type side, we found that participants felt similarly about sharing the same type of data from different devices (e.g., sharing occupancy data from smart door locks and smart thermostats), but their attitudes are more divided when it comes to data types. Participants are generally unwilling to share video, audio, and home occupancy data, while caring less about utility data (e.g., power or heat consumption). The sensitivity of device usage history (status changes or commands), however, varies depending on the device type. The usage history of smart door locks is considered the most sensitive, as it contains the lock status of the door, while smart thermostats' usage history is the least.

Depending on data recipients, data types, and people's original opinions about data sharing, one might prefer different access control mechanisms (e.g., in what situation someone can or cannot access the data, for how long, and so on). We found that in almost all situations, participants would be very upset if someone gained access without explicit approval. Other than explicit approval, whether the data recipients, especially the domestic workers, are on site or not could greatly change one's opinion about data sharing. Remote access for them may be undesirable unless otherwise specified. Device location, on the other hand, could be a crucial contextual factor for roommates' access, as shared devices are likely to be placed in the common area, while everyone's own device will be in their rooms.

To sum up. we make the following important **contributions** in this paper:

- Following the theory of contextual integrity, we conducted a large-scale vignette survey ($N = 1,992$) to gauge people's preferences for interpersonal data sharing in a smart home context.
- We performed an in-depth analysis of the survey results, showing how data recipients, data types, and sharing principles are associated, and how they, together, affect people's access control decisions.
- We summarized a list of design implications for future smart home systems, informing designers and researchers of the potential simplification we can do for future context-adaptive systems.

2 BACKGROUND & RELATED WORK

Privacy in a smart home can be highly contextual [7, 23, 38]. People's privacy preferences can vary between different scenarios, and it is hard for anyone to navigate the sea of contexts that can influence one's privacy attitudes.

In this section, we first introduce the Contextual Integrity framework, a theoretical privacy framework that helps privacy researchers structuralize privacy-related contexts. It assists us in disentangling contextual factors involved in interpersonal data sharing. We then discuss past research on privacy settings designs and contextual access control, and how our work differs from them. A direct comparison between our work and prior studies can be seen in Table 1.

Ref.	CI-based?	*Interpersonal* data recipients	*Data* or *device*?	*Interpersonal* contexts
[3]	✓	10	◑ Data	2
[7]	✓	3	● Both	3
[9]	✓	—	◑ Data	—
[13]	✗	—	◐ Device	—
[23]	✗	6	● Both	8
[38]	✗	—	● Both	—
Ours	✓	11	● Both	11

Table 1: Comparison with prior works. *Non-interpersonal* data recipients (e.g., manufacturers, advertisers) and contexts (e.g., encryption, data storage) are not counted in the table.

2.1 Contextual integrity in smart homes

The Contextual Integrity (CI) framework is one of the most well-established frameworks to contextualize one's privacy attitudes [39]. The CI framework is a normative model "for evaluating the flow of information between agents (individuals and other entities), with a particular emphasis on explaining why certain patterns of flow

provoke public outcry in the name of privacy (and why some do not)" [10]. The CI framework has been widely used to study privacy norms in various scenarios, and the smart home is no exception.

Several studies have used the CI framework to study people's privacy preferences for sharing data in a smart home. For example, prior work has studied people's attitudes towards sharing data with manufacturers, third parties, ISPs, governments, and advertisers [3, 7–9, 38]. Other studies explore privacy attitudes towards personal relationships, but their scope is limited to one or a few particular types of users [7, 35], or a particular device [3]. Therefore, we conduct a study that systematically examines the privacy norms of various types of users, devices, data, and access control mechanisms. Our work compares the norms of various scenarios to understand the similarities and variations across circumstances better. Such findings can guide future smart home designers to develop the system more effectively.

2.2 Smart home privacy research and access control

Understanding people's perceptions of smart-home devices in their daily lives is becoming an increasingly important field of research as the market grows. Previous research found that current smart-home systems frequently fail to meet users' privacy demands and expectations [27, 45]. Some studies thus designed various smart home interfaces and systems to enhance the data transparency in a smart home and enable better privacy settings [14, 16, 36, 44]. Through surveys, interviews, and co-design workshops, researchers find that users have a strong need for data control [12, 49].

Access control is needed to enable users to control how their data can be accessed. Prior work has demonstrated that traditional access control methods often fall short of recognizing interpersonal dynamics and contexts in a home setting [13, 23]. Several research initiatives have thus explored the addition of *contextual factors* to smart homes' access control design [24, 43, 51]. Despite being technically plausible [43], a context-aware access control system can be complicated to use and difficult to capture the desired contexts, resulting in erroneous system behaviors and frustration for the users [24, 51]. For example, Zeng et al. conducted a study among a small sample of homes, consisting of only long-term residents (e.g., immediate family members, couples, and respectful roommates). Given their trusted relationship and frequent device usage, creating detailed access is unnecessary [51]. The reality, however, is much more complicated. Roommates can be strangers met through leasing companies, and one's relationship with their spouses or significant others may turn sour and become less trusted [11]. With the rise of integrated multi-vendor, multi-user home IoT platforms facilitated by vendor agnostic technologies like Matter [15], more people, such as landlords, handypersons, and domestic workers, will also be involved in a smart home scenario. Context-aware access control is still a need in those situations.

A question thus surfaces – how can we make contextual access control systems more usable? Finer granularity can better capture users' intentions, but making it too fine can also easily render the system unusable. Simplification of a context-aware system is thus needed. To understand how one can simplify a context-aware system, we must have a broader understanding of how various factors (e.g., data recipients, data types, sharing principles) affect people's comfort with data sharing and each other, which is the goal of this paper. Some prior studies' focuses are not on interpersonal data sharing [9, 13, 38], and even if they are, their exploration is limited, making it hard to identify similarities and differences between scenarios, which is key to simplification. Abdi et al. studied a wider range of interpersonal data recipients, comparable with our study, but their exploration of data types (only data from voice assistants) and contextual factors are limited, failing to construct a more holistic view of access control for smart homes [3]. Similarly, He et al. explored various data recipients and data types from different devices, but their focus is more on device control than data sharing; they also failed to explore the correlation between contextual factors to data recipients and types [23]. Without understanding the correlation between these design aspects, all contextual factors are equalized and up to users' choice, leading to an over-complicated design. Our work, on the other hand, attempts to understand which contexts, or sharing principles, are more critical to users' decision-making process. These findings can eventually inform a more efficient design for context-aware access control in smart homes.

3 METHODS

Our study mainly involves a vignette about participants' comfort with data sharing in various contexts. Using the Contextual Integrity (CI) framework as the guideline, we vary parameters like data recipients, data types, and sharing principle to gain a more holistic view of the norm of interpersonal data sharing in smart homes. In this section, we detail our application of the CI framework to survey design, the recruitment process, and the analysis methods.

3.1 Survey design

According to the CI framework, there are five parameters that may affect people's perception of privacy, including data senders, data recipients, data types, data subjects, and transmission principle. When designing the survey, we first defined each variable as follows.

3.1.1 Data senders. A *data sender* is the person who makes the decision to share data with other smart-home users, or, in our context, the one who grants other users access to some part of the data. We assume the smart home's primary user is our survey's main data sender. The primary user here is the one who lives with, controls, and manages the smart home devices. It does not matter if they are the homeowner or the tenants, but we do assume they have control over who they would like to share their smart-home data. The participants will enter the survey as a data sender. They will be informed through an imaginary scenario, detailed in Section 3.2.

It is possible that users other than the primary user can also be the data sender. However, assigning participants to different roles in the same survey can easily cause confusion. A between-subject design may solve the problem, but it will create a sample size that is too big for us to possibly collect. Therefore, we only measure participants' perceptions of other users being data senders.

3.1.2 Data recipients. Since our work focuses on interpersonal data sharing, we only consider data recipients who may interact with the primary user of the smart home directly.

We first consulted prior work that studied privacy issues in interpersonal relationships to form a representative list of relationships that may happen in a smart home [3, 7, 23]. We also conducted a brainstorming session to account for potential smart home users that were not previously studied. It led us to 11 different types of data recipients, shown in Table 2.

Conditions	Types	
Data Recipient	Spouse/Significant other [23]	
	Kids (12-year-old) [3, 23]	
	Guests [3, 13]	
	Neighbors [3, 13, 23]	
	Roommates [3, 13, 23]	
	Landlords [50]	
	Handyperson [5, 13]	
	Babysitters [13, 23]	
	Insurance agency representatives	
	Local law enforcement [3, 7]	
	In-home caretakers [13, 35]	
Device and Data Types	Door locks [3, 7, 13, 25, 34]	Home occupancy
		Visitor
		Usage history
	TVs/Streaming devices [34]	Audio clips
		Watching history
		Usage history
	Security cameras [3, 13, 34, 38]	Audio clips
		Home occupancy
		Video clips
		Usage history
	Thermostats [3, 7, 13, 19, 34]	Home occupancy
		Utility usage
		Usage history
	Voice assistants [7, 13, 19, 34]	Audio clips
		Home occupancy
		Usage history
	Lightbulbs [13, 34]	Home occupancy
		Usage history

Table 2: List of data recipients and data types we included in the survey. Full definitions can be found in Appendix C.

3.1.3 Device and data types. Our work considers several data types that are collected based on our selection of IoT devices. We chose seven IoT devices to use in our study based on popularity and controversy (i.e., most often discussed in prior work): smart locks, smart thermostats, smart TVs, smart light bulbs, voice assistants, security cameras, and smart plugs. However, smart plugs were removed from the selection of devices as different privacy implications may occur caused by devices plugged into the outlet, not the smart outlet or plug itself.

Based on the selected devices, we studied the data types produced by these devices by consulting prior work [3, 23] and existing smart home devices. We set up several inclusion criteria when considering what data types to include. We only consider data types directly collected through the given device's sensors. Therefore, account information, such as email addresses, photos, home addresses, and names, is not considered part of the study. In addition, data collected by the device's manufacturer but not shown to the users, such as GPS data from users' smartphones, is also not included in the study to avoid confusion. With these criteria established, all the team members gathered and discussed what data types should be included in the survey until an agreement was reached. It left us with 18 data types from the six selected devices. The full list of data types for each device is shown in Table 2. The definition of each device and data type can be found in Appendix C.

3.1.4 Sharing principles. Our survey assumes that data sharing in smart home systems is governed by access control. [1] Under this assumption, sharing data with someone is equivalent to granting them access to that data. Therefore, we consider sharing principles as contextual factors that can be built into access control components of a smart home system. For example, one's access to data can depend on whether the data captured is from a private or common area, making the device's location a contextual factor for access control.

We consider two types of contexts: system contexts and social contexts. For system contexts, we consulted prior literature on potential contexts one can use for access control [22–24, 38]. To keep the scale of the study manageable and to focus on the ones that are most informative for designing smart homes, we only consider contexts that are practical for today's smart home system, such as notice and choice (either no consent needed or consent on each use), time (a window of visible events), location (access conditioned on device location or accessor location), and content (access to content that either is or is not about the accessor) [22, 24, 38].

For social contexts, we mainly considered two types of social contexts: verbal promises (verbal promise not to share data) and legal bounds (legal promise not to share data) [51]. We acknowledge that it is not a complete list, but social contexts can easily differ from situation to situation. For example, although *purpose of data access* has been proven to be a powerful indicator of people's willingness to share data [1, 7, 38], it is not included in our study, as it is harder to apply to interpersonal relationships, whose needs and purposes are often more versatile than collective entities like third-party services.

The full list of contexts we considered and their definitions can be found in the survey texts we included in Appendix A.

3.2 Survey instrument

Our survey revolves around three factors: the data recipient, the data type, and the sharing principle. These factors allow us to assess privacy—with context—in the smart home setting. We iteratively piloted the survey to assess the timing, fatigue, and quality of responses. The full text of the survey may be found in Appendix A. Here, we briefly describe the survey.

[1] Sharing data through methods bypassing access control (e.g., taking screenshots of a device's usage history and sharing them through emails) is out of the scope of this paper.

The survey started by informing participants that they would be commenting on scenarios involving smart-home devices. Participants were initially selected based on experience with these devices (crowdsourcing platform profiles), confirmed in the survey prior to participant consent (for an overview of these devices, see Table 2). Participants were asked to imagine that they currently owned one of these devices (e.g., a smart door lock) and that this device was rolling out a new feature allowing fine-grained control over the sharing of device information. For example, the participant could allow a handyperson to see who has opened or locked a smart door lock, but not what the passcode to the door is. Each participant was then randomly assigned a relationship and device type when asked about their comfort in sharing data. For example, how comfortable would you be (five-point Likert) when sharing your {viewing history} from your {smart TV} with your {babysitter}?

We then addressed questions related to sharing principles. If a participant said they were uncomfortable or somewhat uncomfortable with the initial scenario, we then asked a series of follow-up questions focusing on how that discomfort might change depending on a different sharing principle. Questions here formed a matrix (five-point Likert) and were grounded on five themes: notice and choice (no consent needed or request consent on each use), time window of use (access restricted to a time period), location (access only when the device is in a certain place or only if the accessor is in the home or only if no one is in the home), content (only when accessed data is about home's occupant or only when access is not about home's occupant), and externally-enforced restricted access (only after a verbal or legal promise to not disclose data). The selection of these conditions is described in the previous section (Section 3.1). We asked this matrix of questions for both a day-to-day scenario and an emergency scenario.

On the other hand, if the participant said they were comfortable or somewhat comfortable, this evidenced a lack of privacy concern for the setting, allowing us to take a slightly different approach. We first asked a similar matrix of questions aimed at sharing principles. These questions were grouped around notice and choice (approval or no approval needed to see data), location (the device is in a private location, accessor is not in the home, no one or someone is in the home), and further sharing (accessor can share data with others). We then looked further at sharing among relationships by asking whether the participant would be comfortable if the accessor shared this information with a set of 11 relationships from Table 2. If the participant said they were neither comfortable nor uncomfortable, we also asked this set of relationship-based questions, in addition to asking the participant (free-text) to describe any scenarios in which they might be uncomfortable with sharing the data in this setting.

We ended the survey with an assessment of mobile privacy preferences via the Mobile Users' Information Privacy Concerns (MUIPC) [48] and demographic questions.

3.3 Recruitment

We recruited participants via Prolific [26], an online crowdsourcing platform, between June and August 2023. Studies have shown researchers prefer Prolific because the data collected is more reliable and of higher quality than that of other platforms, such as Amazon Mechanical Turk and CloudResearch. Based on data quality measures such as honesty, attention, reliability, and comprehension, Prolific consistently delivered the highest quality of data [40, 41]. We estimated the survey to take 10 minutes. Participants were paid $2 for successful completion of the survey. Our university's Institutional Review Board (IRB) has reviewed and approved the study.

After receiving the IRB approval, we use Prolific's pre-screeners to ensure that the participants are at least 18 years old, reside in the U.S., have an approval rate over 95% on Prolific, and own the smart home device that is assigned to them when they enter the survey.

Table 3 shows the demographics of our participants. Our participants are mostly gender-balanced. Over 60% of them are between 25-44. They are more educated than the general public of the U.S. 69.5% reported to have a bachelor's degree or above. The average duration of our study is 10.6 minutes, with a median of 8.5 minutes.

Gender	**Percentage**
Male	51.6%
Female	45.9%
Non-binary	2.2%
Not listed	0.2%
Prefer not to say	0.2%
Age	
18-24	11.6%
25-34	35.6%
35-44	26.5%
45-54	13.7%
55-64	8.4%
65+	4.1%
Prefer not to say	0.2%
Education	
Some high school or less	0.5%
High school diploma or GED	10.4%
Some college, but no degree	19.1%
Bachelor's degree	41.2%
Associate's or technical degree	11.5%
Graduate or professional degree	16.8%
Prefer not to say	0.4%

Table 3: Demographics of the survey participants.

3.4 Analysis

We placed two attention checks in our survey. Anyone who failed both attention checks was automatically removed from the analysis.

We employed a five-point Likert scale to measure participants' comfort in sharing data with someone. During the analysis, we first assigned numeric values, ranging from -2 to 2, to the five scale points. Here, -2 represents "uncomfortable" and 2 represents "comfortable". We then calculated the mean of participants' responses and referred to it as *average comfort score* in the rest of the paper.

In addition, for all Likert-scale questions, we used Kruskal-Wallis tests first to determine if there was a significant difference among various groups. If the test result showed a significant difference

among groups, we then performed pairwise Mann-Whitney U tests with Bonferroni correction to identify which groups were significantly different. For cases where we would like to examine if a correlation between two variables existed, we used Chi-square tests.

We also use logistic regression to model people's comfort with data sharing (normalized). We first used data recipients, data types (of different devices), MUIPC, and demographic information as independent variables. *Neighbors* and *Camera: Video* are chosen as the baseline for data recipients and data types because they receive the lowest average comfort score among all groups. For each demographic category, we picked the largest group as the baseline. After discovering that MUIPC and demographic information have an insignificant impact on the output (Appendix B), we removed them from the model and re-run the logistic regression, with data recipients and data types being the independent variables. We then used the model to rank data recipients and data types based on their contribution to people's comfort in data sharing. For all statistical tests, we set $\alpha = 0.05$.

Although our survey collected some qualitative data, we only performed ad-hoc analysis due to the volume of collected data. Therefore, we did not present qualitative results in this paper except for using several quotes. Our results thus focus on quantitative analyses instead.

3.5 Limitations

Our study has limitations that are typical of user studies. For one, participants on crowdsourcing platforms are typically younger, more educated, and more technologically savvy than the general population [4, 17, 18, 29, 30]. Additionally, participants may have been biased in their responses due to social desirability (e.g., attempting to provide likable answers) or demand effects (i.e., inferring a study purpose when responding) [31, 32]. To reduce these biases, we avoided mentioning "privacy" in the study's introduction and instead phrased the study as a survey on preferences for data sharing. We also used neutral statements throughout the survey instrument (Appendix A), used gender-neutral names for our vignettes, and provided participants with options to select "n/a" when necessary. Participants were filtered for familiarity with smart devices. We saw the trade-off for a limited participant pool as worth the gained ecological validity by knowing participants were familiar with the devices they were commenting on. Our participants were also limited by being in the United States, so our findings may not generalize to other countries with different cultures. For example, some cultures' views on familial relationships are different from people from the US, which may lead to different levels of comfort in smart home data sharing.

The study may also have been impacted by our look at several, but not all, smart-home devices, data types, and transmission principles. Despite our efforts to be comprehensive about realistic smart-home situations, different contexts may provide different outcomes. We leave these to future work.

Finally, similar to all self-reported privacy research, our results may suffer from the privacy paradox—participants' actual behaviors may be inconsistent with their self-reported attitudes [33]. We thus discuss our results with care in Section 6.

4 RESULTS: DATA SHARING PREFERENCES

In this section, we detail our findings about how people's comfort in data sharing changes based on data recipients (RQ1), device types, and data types (RQ2).

4.1 Overview

After collecting responses from 1,992 participants, we found that data recipients influence participants' comfort in data sharing the most, as shown in Figure 1. Over half of the participants are comfortable sharing data with people who live with them, such as spouses or significant others, kids, roommates, and in-home caretakers. They are least comfortable with sharing data with their neighbors.

The effect of devices or data types on people's decision process is less prominent than on data recipients. We rarely observe significant differences in our participants' attitudes among the same type of data collected by different devices, while their opinions differ by different types of data collected by the same device. The result indicates people's privacy attitudes are more influenced by data types than device types, which shows the potential to simplify the data-sharing mechanisms for smart homes.

4.2 Data recipients

We analyzed how participants' comfort regarding data sharing varies across various smart home users, using pairwise Mann-Whitney U tests and logistic regression. The main results of the latter can be found in Table 4. Based on our results about their comfort for data sharing in general, we can roughly divide potential users in a smart home into three categories: *long-term residents, domestic workers, and incidental users.*

4.2.1 Participants are comfortable sharing data with long-term residents. We found that participants are generally comfortable sharing data with spouses (or significant others), children, and roommates. Over 60% of the participants chose "somewhat comfortable" or "comfortable" when asked about sharing data with these three types of users. We also ran a logistic regression for participants' comfort in data sharing, as shown in Table 4. It turns out that these data recipients are the only ones who receive an odds ratio over 10, meaning that the participants are over ten times more comfortable sharing data with them than with neighbors. Given that these types of data recipients stay at the smart home longer than any other user types, we categorized them as *long-term residents*.

Contrary to prior studies [23, 38, 51], which find children are often subject to stricter access control policies, our participants express comfort in sharing smart home data with their kids. Two-thirds (66.7%) of participants whose assigned data recipient is kids expressed comfort with data sharing, comparable to those assigned with roommates (66% expressed comfort, $p = 1.000$). One possible reason is that prior studies focused more on access control for device control than data sharing. The former raises concerns about children misusing devices and causing safety hazards, while the latter is more harmless.

We also found participants are significantly more comfortable sharing data with their spouses than any other parties listed, including roommates ($p = 0.001$), despite us explicitly mentioning

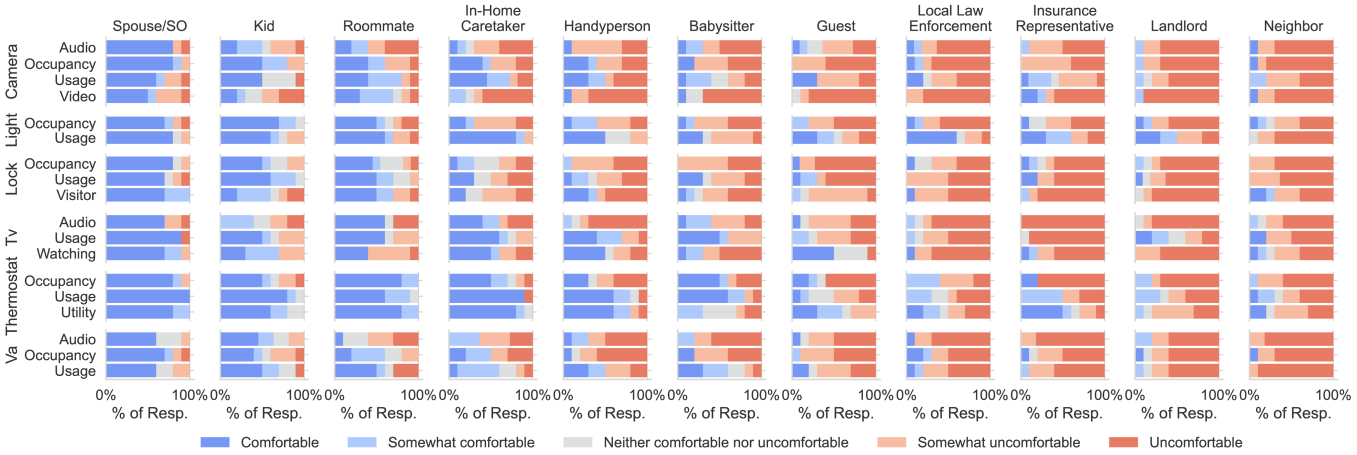


Figure 1: The distribution of participants' comfort of sharing different data with various stakeholders in a smart home. The two-level y-axis details each row's device and data types. Each row shows participants' responses regarding the device and data type listed on the left, while each column shows their responses regarding the recipient type listed on the top.

that the participants share the smart home device with their roommates in our provided scenario. This suggests a preference for data separation even when the smart home devices are equally shared.

4.2.2 Participants are more comfortable sharing data with domestic workers than other incidental users. We ran a logistic model on participants' self-reported comfort in data sharing when being presented with a randomly selected pair of a data recipient and a data type. For the two categorical independent variables, data recipients and data types, we select *neighbors* and *cameras' video data* as the baseline, respectively, because they are the groups with the lowest average comfort score. Table 4 shows the odds ratios, 95% confidence intervals, and the p values for all data recipients and data types. It turns out that participants' data-sharing attitudes towards handyperson, babysitters, and guests are significantly different from those towards neighbors, while such significant differences are not observed among other incidental users (i.e., non-long-term residents), such as local law enforcement, insurance agency representatives, and landlords.

It may be expected that some participants feel comfortable sharing data with guests, as they are from the primary users' social circle. The remaining incidental users are all strangers, or acquaintances at best, but interestingly, participants are generally more comfortable sharing data with domestic workers, such as handypersons and babysitters, than other types of incidental users. One key difference between domestic workers and others may be the fact that they are actively hired by the primary user for help, which provides a potential reason for sharing data with them. Although other incidental users may also have a reason for asking for the data, it is less clear whether offering data without restriction will actually help the primary user. For example, the insurance company may deny the primary users' compensation request based on the given data. A landlord or police may even use the provided data against the primary user, making the latter less comfortable allowing such access.

4.3 Device types vs. data types

Among all smart home devices, participants are most comfortable sharing data from a smart thermostat ($mean$ = 0.412), followed by a smart lightbulb ($mean$ = 0.145). Unsurprisingly, participants feel least comfortable sharing data from a smart security camera ($mean$ = −0.535), closely followed by a voice assistant ($mean$ = −0.512).

It is worth noting that the statistics about a device are largely affected by the type of data the device collected. We found that people are most comfortable sharing utility data (e.g., electricity and heat) and usage history data (e.g., what time the device is on or off), with a mean comfort score of 0.645 and 0.099, respectively. It is also unsurprising that participants are generally uncomfortable sharing video ($mean$ = −0.918) and audio data ($mean$ = −0.724).

4.3.1 Data types matter more than device types. Much prior work has studied the privacy norms on the granularity of devices. Many smart home systems (e.g., SmartThings) also group data by devices. One smart home device, however, can collect different types of data, and different devices may collect data of the same type. This raises questions about whether people's privacy attitudes change across device types or data types.

We found that device types had a limited impact on people's comfort with data sharing. Three types of data in our study are collected by multiple smart home devices – occupancy, audio clips, and usage history (e.g., when the device is on/off, with no video or audio attached when applicable). Among these three data types, we only observed significant differences among usage history collected by different devices, which we will discuss in the next subsection.

On the contrary, for five out of six smart home devices, we observed significant differences among different data types collected by one device. The smart door lock is the only device for which we did not observe a significant difference among various types of collected data (p = 0.690). As expected, people were least comfortable sharing audio or video data than other types of data: 67.2% of our participants expressed discomfort with sharing audio, and 70.9%

	Odds Ratio	95% CI	p
Intercept	0.082	[-3.108, -1.896]	**<0.001**
Recipients (v. Neighbors)			
Landlord	1.091	[-0.448, 0.623]	0.749
Insurance rep.	1.481	[-0.124, 0.909]	0.136
Law enforcement	1.577	[-0.058, 0.970]	0.082
Guests	2.166	[0.272, 1.274]	**0.003**
Babysitters	3.174	[0.662, 1.648]	**<0.001**
Handyperson	3.636	[0.799, 1.782]	**<0.001**
In-home caretakers	6.604	[1.396, 2.379]	**<0.001**
Roommates	10.786	[1.875, 2.882]	**<0.001**
Kids	12.150	[1.989, 3.005]	**<0.001**
Spouses/SOs	27.072	[2.739, 3.858]	**<0.001**
Devices + Data Types (v. Cameras: Video)			
VA: Audio	1.235	[-0.435, 0.858]	0.522
Camera: Audio	1.432	[-0.285, 1.003]	0.274
TV: Audio	1.459	[-0.268, 1.023]	0.251
Lock: Occupancy	1.691	[-0.113, 1.163]	0.107
VA: Occupancy	1.806	[-0.048, 1.230]	0.070
Camera: Occupancy	1.990	[0.052, 1.325]	**0.034**
Lock: Visitor	2.014	[0.064, 1.337]	**0.031**
Lock: Usage	2.315	[0.207, 1.471]	**0.009**
Light: Occupancy	2.407	[0.245, 1.511]	**0.007**
VA: Usage	2.581	[0.316, 1.580]	**0.003**
TV: Watching	2.832	[0.410, 1.672]	**0.001**
Camera: Usage	3.372	[0.588, 1.843]	**<0.001**
Thermostat: Occupancy	3.479	[0.617, 1.877]	**<0.001**
TV: Usage	3.942	[0.742, 2.002]	**<0.001**
Light: Usage	6.601	[1.251, 2.524]	**<0.001**
Thermostat: Usage	7.342	[1.354, 2.633]	**<0.001**
Thermostat: Utility	7.891	[1.424, 2.707]	**<0.001**

Table 4: Logistic regression for participants' comfort in sharing various data with different data recipients. We use neighbors as the baseline for data recipients, and cameras' video data for data types as they receive the lowest average comfort score among all groups. The larger the odds ratios are, the more comfortable participants feel about the listed conditions. The model here omits participants' MUIPC responses and demographics, because they have insignificant impact on the outcome. A full model with all measured variables can be found in Appendix B.

for videos. What intrigues us, however, is that our participants are significantly more open to sharing the usage history of cameras or voice assistants, as long as videos or audio (including transcripts) are not attached.

We also found that participants were often uncomfortable sharing occupancy data. Across all devices, 59.4% of participants expressed discomfort sharing occupancy data with others, which is significantly higher than data types like usage history (44.5%, $p < 0.001$) or utility data (28.2%, $p < 0.001$).

4.3.2 The sensitivity of usage history depends on device types. As mentioned previously, we found that participants' comfort in sharing the usage history of a device is significantly correlated with the given device's type ($p < 0.001$). Among all devices, the participants believe that the usage history of smart door locks is the most sensitive, as it indicates whether or not the door is locked. On the other hand, although smart security cameras are often viewed as the most privacy-sensitive smart home device, participants are pretty neutral about sharing the camera's usage history data (whether the camera is triggered, with no videos attached), resulting in an average comfort score of -0.027. One participant wrote: *"...Just a camera being activated seems not that important unless no one is supposed to be home. A pet could also cause this. It doesn't tell you much otherwise, ..."*

The participants are most comfortable sharing the usage history of a smart thermostat. Compared to the usage history of smart thermostats ($mean = 0.591$), participants are significantly more uncomfortable sharing the usage history of smart door locks ($mean = -0.339$, $p = 0.001$) and voice assistants ($mean = -0.236$, $p = 0.003$). The other two data types show no significant differences when collected by different devices ($p = 0.927$ for audio data, $p = 0.191$ for occupancy).

5 RESULTS: SHARING PRINCIPLES

As previously discussed, the participants are generally uncomfortable sharing smart-home data with others, unless they are long-term residents. Therefore, it is crucial for smart home systems to deploy proper access control mechanisms, not only for device control, but for data access as well. The question is, how should we design the access control system for smart homes? In this section, we discuss contexts and access control mechanisms that may influence people's data-sharing preferences (RQ3).

5.1 Explicit approval is necessary

Whether the primary user has given explicit approval for access is the most critical factor regarding data sharing, as Figure 2 suggests. On average, 51% of all participants who originally expressed uneasiness about data sharing reported that they were at least somewhat comfortable sharing data with someone as long as explicit approval was given. Such attitude changes were especially notable when the data recipient was a kid or a domestic worker. It means that participants acknowledge these parties may need access to the data, but such access must be given explicitly.

Interestingly, for all the participants who originally were comfortable or somewhat comfortable sharing data, 47% of them changed their answer to uncomfortable or somewhat uncomfortable if explicit approval was *not* given. The shift is widespread when the data recipient is a domestic worker, but less so when it comes to the kid. It further shows that explicit approval is completely necessary for allowing domestic workers access to smart home data. For kids, however, it may be more of a personal opinion.

5.2 Contexts that cause discomfort

For participants who originally reported at least somewhat comfortable sharing data with the data recipient, we provide a list of contexts that may cause some concerns. The results show that depending on who the data recipient is and what data type is presented to the participants, the contexts that can cause discomfort also differ.

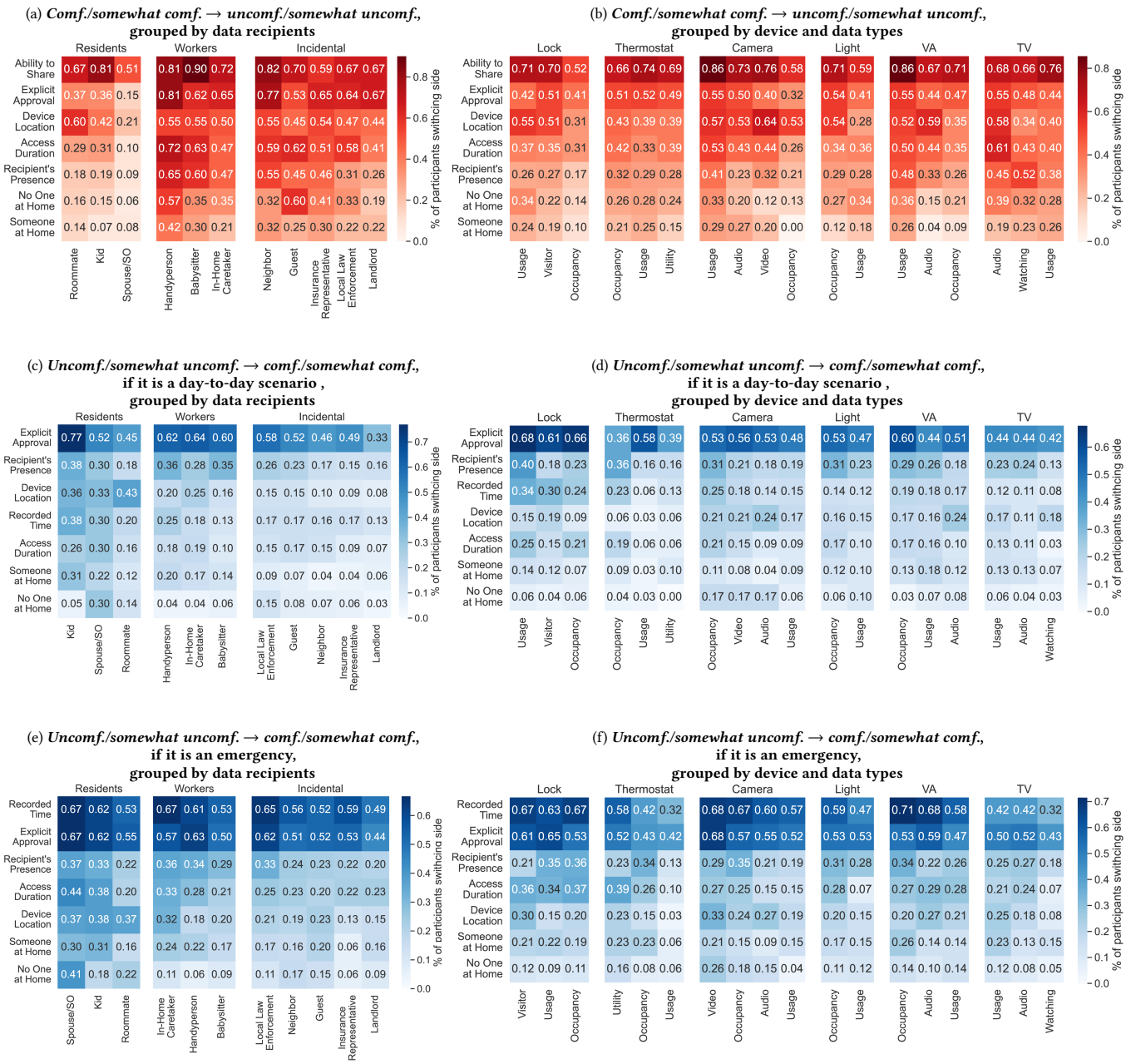


Figure 2: The proportion of participants who switched sides (from "Uncomfortable/Somewhat uncomfortable" to "Comfortable/Somewhat comfortable", and vice versa) when we added more contexts to the original question about their comfort in sharing a particular type of data with a certain person. The contexts are listed on the y-axis of all the heatmaps, and the number in the cells denotes the proportion of people who changed their opinion in a way noted by the title of each figure. The darker the color, the more participants changed their opinion after seeing the contextual factors listed on the y-axis.

5.2.1 Data from private areas are off-limits even for long-term residents. Even though participants are often comfortable sharing data with other residents in the home, even if no explicit approval is given, devices from private areas are likely to be exceptions. This is especially true when the data recipients are the primary user's roommates. As shown in Figure 2a, 60% of the participants, who originally reported they were comfortable sharing data with a roommate, stated that they would be at least somewhat uncomfortable if the device is from a private area. 42% reported the same for kids. Access to data from private areas causes discomfort for more people

than accessing data without explicit approval does. For roommates, accessing data from private areas is significantly more upsetting than accessing data without explicit approval ($p = 0.021$). Such observations are not made in other types of residents. It is possible that people believe long-term residents need to have access to the data for a longer period, making it more likely for unintended data sharing to happen.

5.2.2 Unlimited access duration is unacceptable for domestic workers. Most participants believe that anyone other than long-time residents should only have temporary access to the data. Such belief is most obvious when it comes to handyperson and babysitters. 72% of participants were originally comfortable sharing data with a handyperson, but they would switch sides if there were no time limitations on their access. One thing worth noting is that the number of participants worried about domestic workers accessing data remotely is comparable with unlimited access duration. It indicates that although temporary access is crucial for domestic workers, it does not have to be determined by time. The presence of domestic workers in the home is equally important.

5.2.3 People who trust landlords may not mind them having access remotely. According to Table 4, landlords are among the least welcomed types of users in a smart home system, which is almost comparable to neighbors. That being said, among participants who were queried about *landlords* as the data recipient, 15.1% of them do report that they are at least somewhat comfortable sharing data with the landlords (Figure 1). In those cases, participants are generally okay with landlords accessing their data remotely, as long as the landlord has been given explicit approval and the access is temporary. Only 26% of participants said that remote access would cause them discomfort in this case (Figure 2a).

The same observations also hold for local law enforcement as well. Although only 20.1% of participants are positive about local law enforcement having access to their data, for those who are okay with it, it is mostly acceptable for them to have access remotely as well. Such claims cannot be made regarding domestic workers, neighbors, and guests. These roles are seen as less likely to need access when they are not present in the primary users' homes.

5.3 Contexts that mitigate discomfort

In general, a lot fewer participants changed their opinion if they already felt uncomfortable sharing data with someone. Aside from explicit approval being the most dominating contextual factor, we also found that some contextual factors can sway quite a few participants' attitudes.

5.3.1 Data from common areas can be shared between roommates. Although only 27.2% of our participants reported being uncomfortable or somewhat uncomfortable sharing data with their roommates, it turns out 43% of them would be likely to reconsider their decision as long as the device is from a common area. It echoes our previous conclusion in Section 5.2.1. As a result, device location is the most crucial contextual factor that needs to be considered when sharing the device with a roommate.

5.3.2 Recipient's presence matters more than access duration. We have discussed in Section 5.2.2 how unlimited access duration makes participants less comfortable sharing data with domestic workers. What is the way to mitigate such concern remains unanswered.

Figure 2c shows that other than explicit approval, the recipient's presence is the leading contextual factor that can mitigate the participants' concerns about data sharing with domestic workers and incidental users. When the data recipients were domestic workers, around one-third of the participants changed their opinion to a more positive one once they learned that the data would only be shared when the data recipients were on site. It is likely a result of the expectation that if domestic workers are at one's home, then they are likely working, which indicates a purpose of the need for data access.

5.3.3 No one is home rarely matters for data sharing. In almost all cases, accessing the data when nobody is home did little to mitigate, if not aggravate, people's privacy concerns. The only exception may be the primary users' spouses or significant others. 30% of the participants found them more comfortable sharing data with their spouses or significant others when nobody was home. Interestingly enough, our participants also welcome this contextual factor a little more regarding security cameras than other devices. 17% switched their answer to comfortable or somewhat comfortable when the given device is a camera, while $\leq 10\%$ did so with all other devices. The reason may be that they do not want their spouses or significant others to spy on them.

5.4 Emergency requires different access control mechanisms

If participants reported uncomfortable or somewhat uncomfortable sharing data, we further asked them how they would behave if it were for an emergency (e.g., fire, theft, medical issues). Aligning with previous research [7, 35], we also found that our participants are more likely to allow access during an emergency, as shown in Figure 2e and Figure 2f.

What interests us, however, is that participants would also prefer to grant access in an emergency based on when the data is recorded, besides explicit approval. The need for using data captured time as a contextual factor is not observed when it is a day-to-day scenario, as depicted in Figure 2c and Figure 2b. In hindsight, it makes sense as emergencies are often one-time incidents, which means it would be easy for participants to pinpoint the time when the emergency occurs, and make decisions based on that. A day-to-day scenario, on the other hand, is more likely to have routines or incidents that happen repetitively.

5.5 Data subjects matter

If participants stated that they were uncomfortable sharing data with the given recipient, we further asked them how comfortable they would be if the shared data involved the data recipient. This question is only asked when the data type is video or audio clips, as other data types may not always associate with a data subject (e.g., home occupancy). The results are shown in Figure 3.

In all situations, at least 20% of participants expressed comfort in sharing video and audio data, if the data recipients are involved in it. For data recipients like spouses (or SOs), roommates, and in-home caretakers, over 40% of participants believe it is fair for these data recipients to access the data if they are involved. Interestingly,

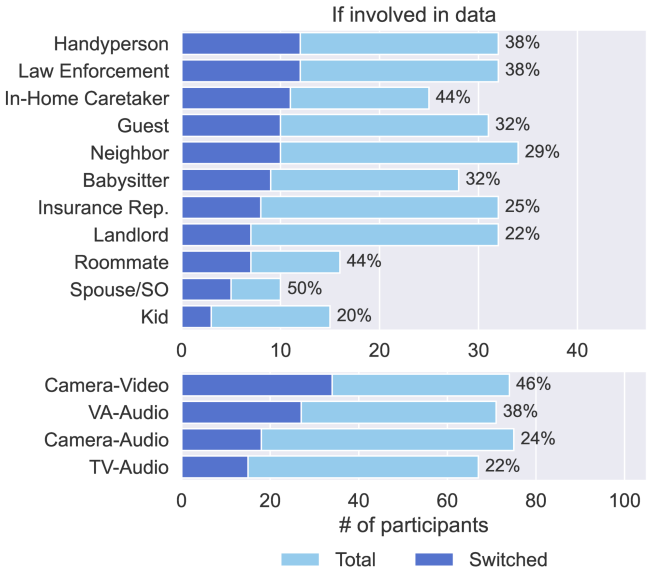


Figure 3: The number of participants who switched from "Uncomfortable/Somewhat uncomfortable" to "Comfortable/-Somewhat comfortable" after knowing the data involves the data recipient.

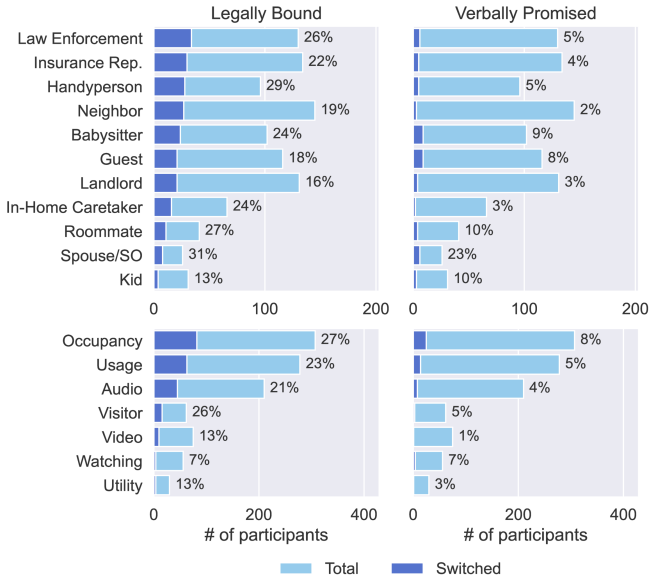


Figure 4: The number of participants who switched from "Uncomfortable/Somewhat uncomfortable" to "Comfortable/-Somewhat comfortable" after knowing the data recipient is legally bound (or verbally promised) not to disclose the data.

our participants do not think the same for children, even though they are also long-term residents. Our guess is that because young children (12 years old in our study) are under the supervision of their parents, the parent's decision about whether the children should have access to the video or audio data has little to do with whether they are present in it.

In addition, we also noticed that 46% of participants are willing to share video data if the data recipient is involved, which is more than those for audio data. We speculate on two potential reasons. One is that participants believe audio data is less sensitive than video data, so it would be less of a concern whether the data recipient is involved. Another reason might be it is less likely for audio data to be accidentally recorded (e.g., a voice assistant often needs a wake-up word to start recording) than video data would be. The users of these devices should be fully aware of the data collection if they actively interact with it, and thus, there is less need for transparency.

5.6 The possibility of delegation

In this section, we discuss how the participants perceive delegation – letting others have the ability to grant access to more people. For participants who initially stated that they were comfortable in data sharing, we further asked how they felt if the data recipients had the ability to share the data with someone else. On the other hand, for participants who initially stated that they were uncomfortable sharing data with the data recipient, we further asked how they felt if the data recipient verbally promised, or was legally bound not to share the data.

5.6.1 Participants do not like the idea of delegation, even for spouses/SOs. Although we anticipated that participants would not like the idea, it still surprised us that it easily became the most influential contextual factor that makes people unwilling to share access in all situations (Figure 2a and Figure 2b). Even for the most trusted type of data recipient, such as spouses or significant others, 51% of participants stated that they would be upset if their spouses or SOs shared the data with someone else.

Some quotes from the participants shed light on the reasoning behind this. It turns out that some participants are worried that their spouse or significant others would share the data without discussing it with them first. For example, when asked about situations where they would feel uncomfortable sharing data with their spouse, one participant wrote: *"Unless I have a reason not to trust Blaire, like she gave the use of the service to someone that shouldn't have it or without mutual agreement, ..."* Another participant, who also shared the same concern, wrote: *"I would not want Blaire to share the data with anyone else without explicit permission."*

5.6.2 Social interactions rarely mitigate participants' concerns. As mentioned in Section 3.1, we considered two types of social interactions: verbal promises and legal bounds. In general, we found that social interactions are less effective in mitigating participants' discomfort with data sharing than many other contextual factors, comparing Figure 4 to Figure 2 and Figure 3. Even if the data recipient is legally bound not to disclose the data, only 22% of the participants, on average, changed their initial opinion across all groups. The number decreases to 5.5% when it comes to verbal promises.

First of all, it seems that no matter whether the data recipients are legally bound not to disclose the data to anyone, or verbally promise it, it means little if the primary user, or the participants here, believe that they do not need the data in the very first place.

Second, compared to verbal promises, legal bounds are much more preferred. For data recipients like local law enforcement, handyperson, and in-home caretakers, over 20% more participants changed their minds when the data recipients were legally prohibited from disclosing the data. Similar to tech support for computers and smartphones, legal responsibility for not disclosing the data may ease peoples' concerns about data sharing, especially when participants may have to give away some data for other benefits (e.g., granting handyperson access to get the device fixed).

6 DISCUSSION

In this section, we talk about the lessons we learned from our results and how these lessons can inform future smart home designs. These lessons can be applied to manufacturers who would like to build a more context-aware access control for their own devices and standard makers building a unified smart home.

6.1 Data sharing and the length of the stay

As discussed in Section 4.2, the 11 types of data recipients we included in the survey can be roughly categorized into long-term residents, domestic workers, and incidental users. Based on participants' responses to each category of users, it seems that the longer one's stay at home is, the more comfortable participants feel about sharing data with them.

The reason behind this could be multi-folded. First of all, longer-term users may have a larger need for access to the data, making their requests for data access more reasonable. Secondly, if someone is not trusted by the primary user, it is unlikely for them to be at their home all the time. It turns out, that how long someone stays at home can be a measurable potential proxy for intention (why they need the access) and trust—the two factors that are recognized as the main motivators for granting access. It not only works for allowing access to longer-term residents, but it might also be an indicator that someone is leaving (e.g., roommates moving out, or breaking up with a former significant other), nudging the primary user to revoke this person's access. Although more research is needed to confirm our hypothesis, it could provide a potentially new perspective on how to set up the default policies for someone and how to make the system adaptive to changes in one's relationships and life.

6.2 Device control vs. data access

Compared to prior research, one interesting thing we noticed is that people's desired access control policies for device control and data access are different. For example, multiple works have discovered that people would often restrict their children's ability to control devices, in case they mess up the system settings or do something unwanted or unsafe [19, 20, 23]. In our study, however, people are generally okay with their kids being able to view the data, except for video data. Similarly, smart thermostats have always been a source of fight over controls [37]. On the other hand, whether someone can view the data on smart thermostats is not something people care very much about. Therefore, viewing and operating the device should be two different categories of access control. A user should be able to specify whether they would like someone to operate the device or view the data for better transparency.

In addition, simply separating the viewing and operating privileges may not be enough. Matter, a new IoT standard [15], has already designed different levels of access in their system, including view, proxy-view, operate, manage, and administer. Each privilege subsumes the capabilities of the prior ones. Although it might be reasonable to assume that someone needs privileges like viewing and operating to manage or administrate devices, comparing our findings with prior works suggests that viewing and operating should be parallel privileges. Someone can allow others to use the device, but not see the history data recorded by it. For example, it is common for guests to use a voice assistant in one's home, but being able to see others' past conversations with the voice assistant may not be acceptable. Therefore, we believe that viewing and operating privileges should be granted or revoked independently.

6.3 The design of temporary access

It is not surprising that temporary access is commonly desired. Many of today's smart home devices or systems have already designed various kinds of temporary access. For example, prior studies have found that smart door locks have four types of users: owner, residents, recurring guests, and temporary guests [22, 25], which is very similar to the categories we made in Section 4.2. Recognizing the fact that access should be temporary by design is merely the first step towards a more privacy-respectful system. Understanding what constitutes "temporary access" and how to design a system that matches people's mental models better is the next step.

In our study, there are mainly two types of contextual factors that decide when someone could or could not have access: access duration and different parties' presence at home. We found that the data recipient's presence at home influences people's decisions more than access duration (Section 5.3.2). It indicates that it may be more intuitive for people to create policies based on the data recipient's presence, than deciding how long or at what time the recipient should have access. Indeed, the reason why people want to specify the duration of the access is likely to limit the access only to the time when the recipient is on site. It also covers cases where a domestic worker or a guest does not have a regular visit schedule, saving the trouble of changing their access's time window or granting access explicitly every time they visit.

That being said, we do not propose to remove access duration as an access control mechanism. The benefit of specifying a time window for access is that it is deterministic. The primary user would know exactly when someone will or will not have access. Granting access based on one's presence does not have such a guarantee. If the data recipient has malicious intentions, they could simply show up at one's place unannounced and gain access. Therefore, although granting access based on one's presence may be more convenient, it can only work in a trusted relationship or for insensitive data.

6.4 Contexts recommendations for explicit approval

Making access control decisions explicitly is often criticized for putting too much burden on the users, asking them for permissions repeatedly, with an overly complex mechanism [2]. Prior works thus have tried to simplify the access control system by automating the decision process for users [9]. Our results, however, show that

people actually desire explicit approval, and can even get upset when none is provided (Section 5.1).

Although the result could be attributed to the privacy paradox, it still reflects people's fear of losing control over their own homes. As noted by Colnago et al., people's concern about losing their autonomy can overpower their desire for convenience [14]. Letting the system make decisions on behalf of the users is thus not a solution, especially when the model's prediction can be complicated to explain.

Therefore, we believe explicit approval is necessary, but we could prioritize certain contexts during the approval process, simplifying the complexity of the interface. For example, long-term residents who live in the household may often need to gain access to data. For these data recipients, a one-time explicit approval may be enough. The main context that needs to be considered is the location of the device, especially when the data recipient is one's roommate. For those one-time or rare visitors, explicit approval could be mandatory every time they visit, and each time the access will be temporary, unless otherwise specified. Given the infrequency of their visit, not much burden will be put on the primary users. The system could also save the previous access configuration, so the future approval process would be simple and quick. Recurring visitors could be trickier to deal with. Depending on whether their visit is regular, one could either choose to use their presence as a trigger for allowing access, or set up recurring time windows for future access.

6.5 Technical solutions vs. social solutions

Technical solutions alone are often insufficient to solve human-centered problems, as people intuitively rely on social norms to make decisions [51]. How to make a smart home system acknowledge social interactions and utilize them for access control would be an interesting research direction.

Although our study did not find a verbal promise useful for people's data-sharing decisions, legal restrictions, on the other hand, actually mitigate some people's concerns. For some types of data recipients we mentioned in the study, such as in-home caretakers, it won't be surprising if they are under some obligations for not sharing information about the care receiver. It would be interesting if such legal promises could be recognized and verified by the smart home system, so that the users can be ensured that the shared data will be handled appropriately. If such legal promises can be obtained and verified by the system, it could become one of the contextual factors for access control in smart homes.

ACKNOWLEDGEMENTS

This research results from a research program at the Institute for Security, Technology, and Society at Dartmouth College, supported by the National Science Foundation under award numbers CNS-1955805, and CNS-1955228. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the sponsors. Any mention of specific companies or products does not imply any endorsement by the authors, by their employers, or by the sponsors.

REFERENCES

[1] Abderrazak Abdaoui, Aiman Erbad, Abdulla Al-Ali, Amr Mohamed, and Mohsen Guizani. 2021. Fuzzy Elliptic Curve Cryptography for Authentication in Internet of Things. *IEEE Internet of Things Journal* 9, 12 (Oct. 2021), 9987–9998. https://doi.org/10.1109/JIOT.2021.3121350

[2] Noura Abdi, Kopo M. Ramokapane, and Jose M. Such. 2019. More than Smart Speakers: Security and Privacy Perceptions of Smart Home Personal Assistants. In *Proceedings of the Fifteenth Symposium on Usable Privacy and Security (SOUPS)*. 451–466.

[3] Noura Abdi, Xiao Zhan, Kopo M. Ramokapane, and Jose Such. 2021. Privacy Norms for Smart Home Personal Assistants. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14. https://doi.org/10.1145/3411764.3445122

[4] Anne M. Turner, Thomas Engelsma, Jean O. Taylor, Rashmi K. Sharma, and George Demiris. 2020. Recruiting older adult participants through crowdsourcing platforms: Mechanical Turk versus Prolific Academic. In *Proc. AMIA*.

[5] Denise Anthony, Carl A. Gunter, Weijia He, Mounib Khanafer, Susan Landau, Ravindra Mangar, and Nathan Reitinger. 2023. The HandyTech's Coming Between 1 and 4: Privacy Opportunities and Challenges for the IoT Handyperson. In *Proceedings of the 22nd Workshop on Privacy in the Electronic Society (WPES'23)*. 129–134.

[6] Apple Inc. 2019. Apple HomeKit. https://developer.apple.com/homekit/

[7] Noah Apthorpe, Yan Shvartzshnaider, Arunesh Mathur, Dillon Reisman, and Nick Feamster. 2018. Discovering Smart Home Internet of Things Privacy Norms Using Contextual Integrity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (Jul. 2018), 1–23. https://doi.org/10.1145/3214262

[8] Noah Apthorpe, Sarah Varghese, and Nick Feamster. [n. d.]. Evaluating the Contextual Integrity of Privacy Regulation: Parents' IoT Toy Privacy Norms Versus COPPA. In *Proceedings of the 28th USENIX Security Symposium* (2019). 123–140.

[9] Natã M. Barbosa, Joon S. Park, Yaxing Yao, and Yang Wang. 2019. "What if?" Predicting Individual Users' Smart Home Privacy Preferences and Their Changes. *Proceedings on Privacy Enhancing Technologies* 2019, 4 (Oct. 2019), 211–231. https://doi.org/10.2478/popets-2019-0066

[10] Adam Barth, Anupam Datta, John C Mitchell, and Helen Nissenbaum. 2006. Privacy and contextual integrity: Framework and applications. In *2006 IEEE symposium on security and privacy (S&P'06)*. IEEE, 15–pp.

[11] Rose Ceccio, Sophie Stephenson, Varun Chadha, Danny Yuxing Huang, and Rahul Chatterjee. 2023. Sneaky Spy Devices and Defective Detectors: The Ecosystem of Intimate Partner Surveillance with Covert Devices. In *Proceedings of the 32nd USENIX Security Symposium*. 123–140.

[12] Chola Chhetri and Vivian Genaro Motti. 2022. User-Centric Privacy Controls for Smart Homes. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 1–36. https://doi.org/10.1145/3555769

[13] Camille Cobb, Sruti Bhagavatula, Kalil Anderson Garrett, Alison Hoffman, Varun Rao, and Lujo Bauer. 2021. "I would have to evaluate their objections": Privacy tensions between smart home device owners and incidental users. *Proceedings on Privacy Enhancing Technologies* 2021, 4 (2021), 54–75. https://doi.org/10.2478/popets-2021-0060

[14] Jessica Colnago, Yuanyuan Feng, Tharangini Palanivel, Sarah Pearman, Megan Ung, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. 2020. Informing the design of a personalized privacy assistant for the internet of things. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[15] Connectivity Standards Alliance. 2022. Matter. https://csa-iot.org/all-solutions/matter/

[16] Anupam Das, Martin Degeling, Daniel Smullen, and Norman Sadeh. 2018. Personalized privacy assistants for the internet of things: Providing users with notice and choice. *IEEE Pervasive Computing* 17, 3 (2018), 35–46.

[17] Djellel Difallah, Elena Filatova, and Panos Ipeirotis. 2018. Demographics and dynamics of Mechanical Turk workers. In *Proc. WSDM*.

[18] 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 *Proc. IEEE S&P*.

[19] Radhika Garg and Christopher Moreno. 2019. Understanding Motivators, Constraints, and Practices of Sharing Internet of Things. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (Jun. 2019). https://doi.org/10.1145/3328915

[20] Christine Geeng and Franziska Roesner. [n. d.]. Who's in Control?: Interactions in Multi-User Smart Homes. ([n. d.]), 1–13. https://doi.org/10.1145/3290605.3300498

[21] Google Inc. [n. d.]. Google Home. https://home.google.com/welcome/

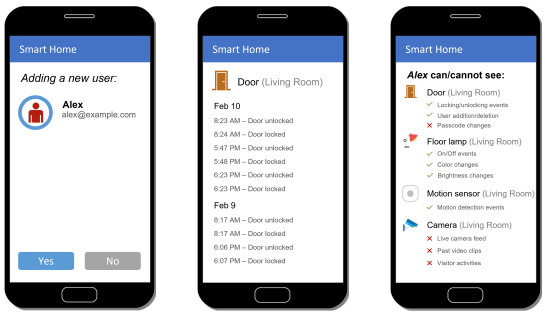
[22] Hussein Hazazi and Mohamed Shehab. 2023. Exploring the Usability, Security, and Privacy of Smart Locks from the Perspective of the End User. In *Proceedings of the Nineteenth Symposium on Usable Privacy and Security*. 559–577. https://www.usenix.org/conference/soups2023/presentation/hazazi

[23] Weijia He, Maximilian Golla, Roshni Padhi, Jordan Ofek, Markus Dürmuth, Earlence Fernandes, and Blase Ur. 2018. Rethinking Access Control and Authentication for the Home Internet of Things (IoT). In *27th USENIX Security Symposium*. 255–272. https://www.usenix.org/conference/usenixsecurity18/presentation/he
[24] Weijia He, Valerie Zhao, Olivia Morkved, Sabeeka Siddiqui, Earlence Fernandes, Josiah Hester, and Blase Ur. 2021. SoK: Context Sensing for Access Control in the Adversarial Home IoT. In *2021 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, 37–53. https://doi.org/10.1109/EuroSP51992.2021.00014
[25] Grant Ho, Derek Leung, Pratyush Mishra, Ashkan Hosseini, Dawn Song, and David Wagner. 2016. Smart Locks: Lessons for Securing Commodity Internet of Things Devices. In *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*. 461–472. https://doi.org/10.1145/2897845.2897886
[26] Prolific Inc. [n. d.]. Prolific. https://app.prolific.co/
[27] Timo Jakobi, Corinna Ogonowski, Nico Castelli, Gunnar Stevens, and Volker Wulf. 2017. The catch (es) with smart home: Experiences of a living lab field study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 1620–1633.
[28] Timo Jakobi, Gunnar Stevens, Nico Castelli, Corinna Ogonowski, Florian Schaub, Nils Vindice, Dave Randall, Peter Tolmie, and Volker Wulf. 2018. Evolving Needs in IoT Control and Accountability. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 4 (Dec. 2018), 1–28. https://doi.org/10.1145/3287049
[29] 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 *Proc. SOUPS*.
[30] Joel Ross, Lilly Irani, M. Six Silberman, Andrew Zaldivar, and Bill Tomlinson. 2010. Who are the Turkers? Worker demographics in Amazon Mechanical Turk. In *Proc. CHI EA*.
[31] Jon A Krosnick, Sowmya Narayan, and Wendy R Smith. 1996. Satisficing in surveys: Initial evidence. *New Directions for Evaluation* 1996, 70 (1996), 29–44.
[32] Jonathan Mummolo and Erik Peterson. 2019. Demand effects in survey experiments: An empirical assessment. *American Political Science Review* 113, 2 (2019), 517–529.
[33] Bernard Lubin and Roger L. Harrison. 1964. Predicting Small Group Behavior with the Self-Disclosure Inventory. *Psychological Reports* 15, 1 (Aug. 1964), 77–78. https://doi.org/10.2466/pr0.1964.15.1.77
[34] Shrirang Mare, Franziska Roesner, and Tadayoshi Kohno. 2020. Smart Devices in Airbnbs: Considering Privacy and Security for both Guests and Hosts. *Proceedings on Privacy Enhancing Technologies* 2020, 2 (Apr. 2020), 436–458. https://doi.org/10.2478/popets-2020-0035
[35] Karola Marky, Alexandra Voit, Alina Stöver, Kai Kunze, Svenja Schröder, and Max Mühlhäuser. 2020. "I Don't Know How to Protect Myself": Understanding Privacy Perceptions Resulting from the Presence of Bystanders in Smart Environments. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction*. 1–11.
[36] Karola Marky, Verena Zimmermann, Alina Stöver, Philipp Hoffmann, Kai Kunze, and Max Mühlhäuser. 2020. All in one! user perceptions on centralized iot privacy settings. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.
[37] Dana McKay and Charlynn Miller. [n. d.]. Standing in the Way of Control: A Call to Action to Prevent Abuse through Better Design of Smart Technologies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021-05-07) *(CHI '21)*. Association for Computing Machinery, 1–14. https://doi.org/10.1145/3411764.3445114
[38] Pardis E Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Cranor, and Norman Sadeh. 2017. Privacy Expectations and Preferences in an IoT World. In *Symposium on Usable Privacy and Security (SOUPS)*. USENIX, 399–412.
[39] Helen Nissenbaum. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.
[40] Stefan Palan and Christian Schitter. 2018. Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17, CSCW2 (2018), 22–27.
[41] Eyal Peer, David Rothschild, Andrew Gordan, Zak Evernden, and Damer Ekaterina. 2022. Data quality of platforms and panels for online behavioral research. In *Behavior Research Methods*. 1643–1662.
[42] Samsung Inc. [n. d.]. Samsung SmartThings. https://www.smartthings.com/
[43] Roei Schuster, Vitaly Shmatikov, and Eran Tromer. 2018. Situational Access Control in the Internet of Things. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, Vol. 18. ACM, New York, NY, USA, 1056–1073. https://doi.org/10.1145/3243734.3243817
[44] William Seymour, Martin J Kraemer, Reuben Binns, and Max Van Kleek. 2020. Informing the design of privacy-empowering tools for the connected home. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
[45] Vandit Sharma and Mainack Mondal. 2022. Understanding and Improving Usability of Data Dashboards for Simplified Privacy Control of Voice Assistant Data. In *31st USENIX Security Symposium*. 3379–3395. https://www.usenix.org/conference/usenixsecurity22/presentation/sharma-vandit
[46] Sophie Stephenson, Majed Almansoori, Pardis Emami-Naeini, Danny Yuxing Huang, and Rahul Chatterjee. 2023. Abuse Vectors: A Framework for Conceptualizing IoT-Enabled Interpersonal Abuse. In *Proceedings of the 32nd USENIX Security Symposium*.
[47] Madiha Tabassum and Heather Lipford. [n. d.]. Exploring Privacy Implications of Awareness and Control Mechanisms in Smart Home Devices. 2023, 1 ([n. d.]), 571–588. https://doi.org/10.56553/popets-2023-0033
[48] Heng Xu, Sumeet Gupta, Mary Beth Rosson, and John M Carroll. [n. d.]. Measuring mobile users' concerns for information privacy. In *Proceedings of the Thirty Third International Conference on Information Systems (ICIS '12)*.
[49] Yaxing Yao, Justin Reed Basdeo, Smirity Kaushik, and Yang Wang. 2019. Defending my castle: A co-design study of privacy mechanisms for smart homes. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–12.
[50] Eric Zeng, Shrirang Mare, Franziska Roesner, Santa Clara, Eric Zeng, Shrirang Mare, and Franziska Roesner. 2017. End User Security and Privacy Concerns with Smart Homes. In *Proceedings of the 13th Symposium on Usable Privacy and Security*.
[51] Eric Zeng and Franziska Roesner. 2019. Understanding and Improving Security and Privacy in Multi-User Smart Homes: A Design Exploration and In-Home User Study. In *USENIX Security Symposium*. 159–176. https://www.usenix.org/conference/usenixsecurity19/presentation/zeng
[52] Serena Zheng, Noah Apthorpe, Marshini Chetty, and Nick Feamster. 2018. User Perceptions of Smart Home IoT Privacy. *Proceedings of the ACM on Human-Computer Interaction* (2018). https://doi.org/10.1145/3274469

A SURVEY TEXT

[validate device, consent, introduction]

Imagine you own some smart home devices, and have an app (Smart Home app) installed on your smartphone that helps you control and monitor all your smart home devices. You could add others through the app. Once a user installs the Smart Home app on their smartphone and is added to your home, they can see all the device activities through the app. Here, for the demonstration purpose, let's assume you want to add a new user, Alex, to the Smart Home app. As shown in the figures below, you can invite Alex to your smart home through the Smart Home app. After being added, Alex needs to install the Smart Home app on their smartphone and accept the invitation. Upon acceptance, Alex can now see all the activities that happened on these smart home devices through the Smart Home app. You can safely assume that the Smart Home app is the only way for Alex to see past events that happened on these devices.



(a) (b) (c)

Figure 5: Images shown the participants in explanation of fine-grained access control

The app recently rolled out a new feature that enables more fine-grained control over what activities other users can see. For example, as shown in the following figure, Alex can see when the smart door is locked or unlocked, and who has gained or lost access to the smart door. However, Alex cannot see if the passcode on the door has been changed or not.

Now, please imagine you have added the following user, {Name}, to the Smart Home App. {Name} is your {relationship}. Assume you have a {device}. This device is shared between you two. The {device} can be unlocked by your smartphone or other Internet-connected devices (based on GPS location), or by typing in a pin code.

In a day-to-day scenario, how comfortable will you be if {Name} can see all the history of the following data captured by the {device}? {device-type} [`e.g., how comfortable would you be (five-point Likert: comfortable, somewhat comfortable, neither, somewhat uncomfortable, uncomfortable, n/a) when sharing your {viewing history} from your {smart TV} with your {babysitter}?`]

uncomfortable or somewhat uncomfortable selected

In a day-today-scenario, how comfortable would you be if the above data ({device-type}) is only shared when the following conditions are met? Please consider each of the conditions independently. NOTE: If any of the following conditions does not make sense to you, please select the last option, "This condition does not make sense."

Five-point Likert: comfortable, somewhat comfortable, neither, somewhat uncomfortable, uncomfortable, n/a

- Every time {Name} wants to see the data, they must request your approval (effective until they close the app)
- {Name} can only see the data that has been recorded during a certain period of time (e.g., from Feb 1 to Feb 7)
- {Name} can only have access to the data for a certain period of time (e.g., for the next three days)
- {Name} can only see the data if the device is not from a private area (e.g., bedroom or bathroom)
- {Name} can only see the data if they are currently in your home.
- {Name} can only see the data if no one is home.
- {Name} can only see the data if someone is home.
- The data involves or is about {Name}
- The data does not involve or is not about {Name}
- {Name} verbally promises that they will not disclose the data they see
- {Name} is legally bound not to disclose the data they see (please assume there exists such kind of laws)
- Please select "Somewhat uncomfortable" for this statement.

`[repeated for emergency scenario, with additional question "{Name} can only see the data recorded during the emergency"]`

In addition to the previous question, are there any other circumstances where you might allow {Name} to have access to this data ({device-type})? If so, please elaborate. If not, please explain why. `[free-text]`

comfortable or somewhat comfortable selected

How comfortable would you be if the above data ({device-type}) is shared in the following scenarios? NOTE: If any of the following circumstances does not make sense to you, please select the last option, "This circumstances does not make sense."

Five-point Likert: comfortable, somewhat comfortable, neither, somewhat uncomfortable, uncomfortable, n/a

- {Name} can see the data without your explicit approval
- {Name} can see all the data whenever they want
- {Name} can see the data even if the device is from a private area (e.g., bedroom or bathroom)
- {Name} can see the data even if they are not currently in your home.
- {Name} can see the data when no one is home.
- {Name} can see the data when someone is home.
- {Name} has the ability to share the data with others.
- Please select "Somewhat uncomfortable" for this statement.

In addition to the previous question, are there any other circumstances where you might allow {Name} to have access to this data ({device-type})? If so, please elaborate. If not, please explain why. `[free-text]`

The current version of the app allows users to share data with other people. Assume that you have shared the following data to {Name}. {device-type} recorded by the {device}. How acceptable would it be if {Name} share this data with the following users, if they exist? NOTE: If you believe that some users in the following questions already know the information, or the overall scenario does not make sense, you could choose an answer under the "Cannot Choose" column.

Five-point Likert: acceptable, somewhat acceptable, neutral, somewhat unacceptable, unacceptable, n/a

- Your spouse/significant other who lives with you
- Your 12-year-old kid who lives with you
- Your guest
- Your neighbor
- Your roommate
- Your landlord
- A handy person
- Law enforcement
- Your babysitter
- Your insurance agency's representative
- Your in-home caretaker

neutral selected

In what circumstances would you allow {Name} to have access to the aforementioned data ({device-type} recorded by the {device})? Please list at least one scenario. You could add more by clicking the button below. `[free-text]`

`[repeat relationship questions (11) from comfortable-somewhat comfortable]`

n/a selected

You have stated that you believe the above scenario does not make sense. Would you mind specifying why you think the given scenario does not make sense?

`[MUIPC]` In this section, you are going to see several statements and questions. There is not correct or wrong answers, so please answer them based on your own experiences. Also, these statements have nothing to do with the previous scenarios, so please answer them independently.

`[same for all questions]` Seven-point Likert: strongly disagree, disagree, somewhat disagree, neither, somewhat agree, agree, strongly agree

To which degree do you agree with the following statements? Please select the one that best represents your opinion.

- I believe that the location of my mobile device is monitored at least part of the time.
- I am concerned that mobile apps are collecting too much information about me.
- I am concerned that mobile apps may monitor my activities on my mobile device.

To which degree do you agree with the following statements? Please select the one that best represents your opinion.

- I feel that as a result of my using mobile apps, others know about me more than I am comfortable with.
- I believe that as a result of my using mobile apps, information about me that I consider private is now more readily available to others than I would want.
- I feel that as a result of my using mobile apps, information about me is out there that, if used, will invade my privacy.

To which degree do you agree with the following statements? Please select the one that best represents your opinion.

- I am concerned that mobile apps may use my personal information for other purposes without notifying me or getting my authorization.
- When I give personal information to use mobile apps, I am concerned that apps may use my information for other purposes.
- I am concerned that mobile apps may share my personal information with other entities without getting my authorization.

How often have you personally experienced incidents whereby your personal information was used by some company or e-commerce web site without your authorization?

- Never (0% of the time)
- Rarely (around 10% of the time)
- Sometimes (around 30% of the time)
- About half the time (around 50% of the time)
- Frequently (around 70% of the time)
- Usually (around 90% of the time)
- Always (100% of the time)

How much have you heard or read during the last year about the use and potential misuse of the information collected from the Internet?

- None at all
- A little
- Some
- A moderate amount
- Somewhat much
- A lot
- A great deal

How often have you personally been the victim of what you felt was an improper invasion of privacy?

- Never (0% of the time)
- Rarely (around 10% of the time)
- Sometimes (around 30% of the time)
- About half the time (around 50% of the time)
- Frequently (around 70% of the time)
- Usually (around 90% of the time)
- Always (100% of the time)

To which degree do you agree with the following statements? Please select the one that best represents your opinion.

- I am likely to disclose my personal information to use mobile apps in the next 12 months.
- I predict I would use mobile apps in the next 12 months.
- I intend to use mobile apps in the next 12 months.

```
[demographic questions]
```

B ADDITIONAL STATISTICS

	Coef.	Std. err.	Odds Ratio	95% CI	p
Intercept	-1.068	0.492	0.344	[-2.033,-0.104]	**0.030**
Recipients (v. Neighbors)					
Landlords	-0.006	0.282	0.994	[-0.558, 0.547]	0.983
Insurance rep.	0.394	0.271	1.483	[-0.138, 0.925]	0.146
Law enforcement	0.515	0.269	1.673	[-0.012, 1.042]	0.056
Guests	0.774	0.264	2.168	[0.257, 1.291]	**0.003**
Babysitters	1.145	0.258	3.142	[0.640, 1.650]	**<0.001**
Handyperson	1.348	0.259	3.849	[0.839, 1.856]	**<0.001**
In-home caretakers	1.932	0.259	6.902	[1.425, 2.439]	**<0.001**
Roommates	2.435	0.265	11.419	[1.917, 2.954]	**<0.001**
Kids	2.533	0.266	12.594	[2.013, 3.054]	**<0.001**
Spouses/SOs	3.437	0.294	31.088	[2.862, 4.012]	**<0.001**
Devices + Data Types (v. Cameras: Video)					
VA: Audio	0.172	0.339	1.188	[-0.492, 0.837]	0.611
Camera: Audio	0.274	0.338	1.315	[-0.388, 0.935]	0.417
TV: Audio	0.394	0.337	1.483	[-0.266, 1.054]	0.242
Lock: Occupancy	0.488	0.334	1.629	[-0.167, 1.142]	0.144
VA: Occupancy	0.607	0.337	1.834	[-0.053, 1.266]	0.071
Camera: Occupancy	0.732	0.333	2.080	[0.079, 1.385]	**0.028**
Lock: Visitor	0.737	0.335	2.089	[0.081, 1.393]	**0.028**
Lock: Usage	0.817	0.331	2.264	[0.168, 1.466]	**0.014**
Light: Occupancy	0.921	0.332	2.511	[0.270, 1.571]	**0.006**
VA: Usage	0.964	0.332	2.623	[0.313, 1.616]	**0.004**
TV: Watching	1.034	0.330	2.812	[0.388, 1.680]	**0.002**
Camera: Usage	1.171	0.325	3.224	[0.533, 1.808]	**<0.001**
Thermostat: Occupancy	1.239	0.331	3.452	[0.590, 1.888]	**<0.001**
TV: Usage	1.404	0.331	4.070	[0.755, 2.052]	**<0.001**
Light: Usage	1.915	0.334	6.787	[1.261, 2.569]	**<0.001**
Thermostat: Usage	1.926	0.337	6.863	[1.266, 2.586]	**<0.001**
Thermostat: Utility	2.032	0.337	7.629	[1.372, 2.692]	**<0.001**
Demographics: Gender (v. Male)					
Transgender/Trans woman	-1.462	1.546	0.232	[-4.492, 1.569]	0.345
Prefer not to say	-0.271	1.734	0.762	[-3.670, 3.127]	0.876
Not listed	-0.034	1.260	0.967	[-2.504, 2.436]	0.979
Non-binary	-0.011	0.365	0.989	[-0.726, 0.704]	0.976
Female	0.065	0.110	1.067	[-0.150, 0.281]	0.553
Transgender/Trans man	0.129	0.865	1.138	[-1.566, 1.825]	0.881

Table 5: Full logistic regression results for participants' comfort in sharing various data with different data recipients. The baseline of all categorical data is included in the parentheses. (Continued in Table 6)

	Coef.	Std. err.	Odds Ratio	95% CI	p
Demographics: Age (v. 25-34)					
18-24	-0.145	0.188	0.865	[-0.513, 0.222]	0.439
55-64	-0.045	0.208	0.956	[-0.452, 0.363]	0.830
35-44	-0.001	0.138	0.999	[-0.271, 0.269]	0.993
45-54	0.047	0.172	1.048	[-0.289, 0.384]	0.783
65+	0.523	0.285	1.687	[-0.035, 1.081]	0.066
Prefer not to say	2.058	2.671	7.831	[-3.176, 7.292]	0.441
Demographics: Education (v. Bachelor's degree)					
Prefer not to say	-1.660	2.098	0.190	[-5.772, 2.451]	0.429
Graduate degree	-0.118	0.157	0.889	[-0.425, 0.189]	0.451
Some college	-0.084	0.151	0.919	[-0.381, 0.213]	0.579
HS/GED	-0.070	0.187	0.933	[-0.437, 0.297]	0.709
Associate degree	-0.057	0.180	0.945	[-0.409, 0.296]	0.753
Some high school or less	0.116	0.773	1.123	[-1.399, 1.632]	0.881
MUIPC: Perceived Intrusion					
Measure 3	-0.202	0.081	0.817	[-0.360, -0.043]	**0.013**
Measure 1	-0.087	0.069	0.917	[-0.223, 0.048]	0.207
Measure 2	0.168	0.083	1.183	[0.006, 0.330]	**0.043**
MUIPC: Secondary Use of Personal Information					
Measure 2	-0.165	0.093	0.848	[-0.347, 0.017]	0.076
Measure 3	-0.022	0.098	0.978	[-0.214, 0.171]	0.824
Measure 1	0.067	0.094	1.070	[-0.116, 0.251]	0.472
MUIPC: Perceived Surveillance					
Measure 2	-0.121	0.078	0.886	[-0.274, 0.033]	0.123
Measure 3	0.017	0.084	1.017	[-0.146, 0.181]	0.836
Measure 1	0.020	0.050	1.021	[-0.078, 0.119]	0.683
MUIPC: Behavioral Intention					
Measure 2	-0.047	0.106	0.954	[-0.254, 0.160]	0.659
Measure 3	-0.008	0.109	0.992	[-0.222, 0.206]	0.943
Measure 1	0.082	0.041	1.086	[0.002, 0.163]	**0.046**
MUIPC: Prior Privacy Experience					
Measure 2	-0.037	0.040	0.963	[-0.115, 0.040]	0.347
Measure 3	-0.005	0.067	0.995	[-0.137, 0.128]	0.945
Measure 1	0.058	0.061	1.060	[-0.061, 0.178]	0.337

Table 6: The continuation of Table 5.

C DEFINITIONS

Data Recipient	Name	Description
Spouse	Blaire	They live with you. Although sometimes you may argue with each other, you two share a very stable relationship.
Kid	Finley	They are your 12-year-old kid who lives with you. They are generally well-behaved, but they are about to be a teenager soon.
Guest	Hayden	They are from your social circle. You talk with each other often, and you have invited them to your home several times.
Neighbor	Riley	They live next door to you. You met them and exchanged greetings with them several times before. They seem nice, but you don't know them very well.
Roommate	Madison	They live with you in a 2 bedroom 1 bathroom apartment for several months now and share most of the appliances with you, except for the ones in your respective bedrooms. They seem to be nice so far and are generally respectful of your boundaries.
Landlord	Robin	They do not live with you. You mainly just talk when the rent is due or when something goes wrong in the apartment. They usually would let you know before they drop by.
Handyperson	Reese	Whenever some of your smart home devices are not working, you will contact a handyperson company to fix the problem. Reese is assigned to you this time. Although you have never met them, they are from a company you have used before. They showed you their work ID when they showed up.
Babysitter	Rowan	You hired them three months ago. They worked at your home for around 10 hours per week. You are satisfied with their work so far.
Insurance representative	Sydney	You don't really know them personally, but they are the person that you will contact if you need insurance services. You have talked to them once or twice before.
Local law enforcement	Jayden	You don't really know them, but if someone calls 911, they are the one that will be sent.
In-home caretaker	Avery	Your parent, who currently lives with you, has a condition that requires an in-home caretaker, so you hired Avery. They have lived with you for three months now, and both you and your parent are satisfied with their work so far.

Table 7: Definitions of data recipients used in our survey. The second column is a gender-neutral name we gave to the date recipients for easier reference later in the survey.

Device	Data Type	Data Type Description
Smart Door Lock	Usage	Locked/unlocked status
	Visitor	Visitor activity
	Occupancy	Occupancy of the home (whether there are people at home)
Smart TV	Audio	Audio clips (from a built-in voice assistant)
	Usage	What time the smart TV is used
	Watching	Watching history
Smart Lightbulbs	Usage	What time the smart lightbulb is used (e.g., on/off, changing color, etc.)
	Occupancy	Whether there are people at home
Smart Security Cameras	Video	Video clips
	Audio	Audio clips
	Usage	What time the camera is triggered (only time, no video)
	Occupancy	Occupancy of the home (whether there are people at home)
Smart Thermostat	Usage	What time the thermostat is used (e.g., on/off, target temperature changes)
	Occupancy	Occupancy of the home (whether there are people at home)
	Utility	Utility usage (e.g., electricity/heat)
Voice Assistant	Audio	Audio clips
	Usage	What time the voice assistant is used (e.g., playing music, controlling other devices, but no audio or transcript)
	Occupancy	Occupancy of the home (whether there are people at home)

Table 8: Devices and their data types and descriptions.

Device	Device Description
Smart Door Lock	The smart door lock allows you to control who enters your home without the need for a physical key. One can unlock the door by your smartphone or smartwatch, or by typing a pin code.
Smart TV	The smart TV allows you to access a variety of online content and streaming services. It also contains a built-in voice assistant and a browser.
Smart lightbulb	The smart lightbulbs allow you to turn lights on/off, dim, change colors and schedules from your phone.
Smart security camera	The smart security camera allows you to monitor activities inside or outside your home. You can view recordings live or afterward.
Smart thermostat	The smart thermostat allows you to control and monitor the temperature of your home remotely. It also can learn your schedule (e.g., whether you are home or not) and adjust the temperature accordingly to save energy.
Voice assistant	The voice assistant is a device that can answer questions, search information, and control smart home devices using voice commands. It can also identify different users' voices and provide services accordingly.

Table 9: Devices used in the survey and our descriptions of them.