Exploring How Privacy and Security Factor into IoT Device Purchase Behavior

Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Henry Dixon
Carnegie Mellon University
E-mails: {pardis, yuvrajagarwal, lorrie, hdixon}@cmu.edu

ABSTRACT
Despite growing concerns about security and privacy of Internet of Things (IoT) devices, consumers generally do not have access to security and privacy information when purchasing these devices. We interviewed 24 participants about IoT devices they purchased. While most had not considered privacy and security prior to purchase, they reported becoming concerned later due to media reports, concerns of friends, or observing unexpected device behavior. Those who sought privacy and security information before purchase, reported that it was difficult or impossible to find. We asked interviewees to rank factors they would consider when purchasing IoT devices; after features and price, privacy and security were ranked among the most important. Finally, we showed interviewees our prototype privacy and security label. Almost all found it to be accessible and useful, encouraging them to incorporate privacy and security in their IoT purchase decisions.

CCS CONCEPTS
• Security and privacy → Usability in security and privacy;

KEYWORDS
Purchasing Behaviors, Comparison Shopping, Privacy, Security, Internet of Things (IoT)

1 INTRODUCTION
While sales of IoT devices are skyrocketing [33], consumers are concerned about the privacy and security of their devices. Surveys have found that privacy is among the biggest concerns consumers have about IoT devices and that people want to have control over the personal information these devices collect [10, 46]. However, there is little information available for consumers who wish to seek out secure and private devices.

Regulators around the world are calling on IoT device manufacturers to implement security safeguards and provide information about device security and privacy. Some have called for standardized IoT product labels that would highlight privacy and security practices [12, 19, 30, 49, 54, 58]. Although these policy reports and proposed legislation advocate for IoT labels, they do not propose specific label designs.

Labels are used in numerous applications such as the nutrition facts label for foods [23], fuel economy and environment label for cars [22], European Union (EU) energy label for office appliances [73], Power Content Label (PCL) for electricity [11], EnergyGuide label for home appliances [27], and Lighting Facts label for light bulbs [26]. Researchers have found that standardized labels are a promising approach for informing consumers about privacy: privacy “nutrition labels” on websites [43], privacy meters in search engines [9], and a “privacy facts” checklist in an app store [44] have been shown to impact study participant decision making. However, labels proposed in this prior work were not designed for IoT devices. In addition, those labels focused only on privacy, not security.

We conducted in-depth semi-structured interviews with 24 participants who had purchased at least one IoT device (smart home device or wearable). We explored interviewees’ understanding of privacy and security issues associated with IoT devices and factors they considered when purchasing their device. At the end of each interview, we displayed prototype IoT security and privacy labels that we developed, and discussed them with interviewees. Finally, we conducted a 200-participant Mechanical Turk (MTurk) survey to probe the influence of privacy and security information when making IoT purchase decisions.

We found that about half of our interviewees had limited and often incorrect knowledge about privacy and security
and that this impacted their ability to make informed privacy and security decisions. In addition, most of our interviewees had not considered privacy and security before purchasing IoT devices, but reported being concerned after the purchase. Those who were concerned about privacy and security at the pre-purchase evaluation stage reported difficulty finding useful information about device privacy and security.

We found that security and privacy are among the factors that people would consider in their future IoT device purchase decisions. Survey participants reported that security and privacy would have significantly more influence on their decisions to purchase a smart security camera than a smart thermostat or toothbrush, likely due to their perceptions of the sensitivity of the data those devices collect. Almost all interviewees acknowledged the importance of knowing privacy and security information related to IoT devices before making purchase decisions and said they would pay a small premium for such information to be provided, especially when purchasing a device that they perceive to collect more sensitive information.

Almost all interviewees found our prototype labels easy to understand and able to provide information they would consider in a purchase decision. We found that interviewees often focused on privacy and security choices, expert ratings, purpose of data collection, and the convenience of security mechanisms. From our findings, we distill recommendations for the design of privacy and security labels that enable consumers to make informed IoT device purchase decisions. Our findings on consumers’ interest in IoT nutrition labels and ways to make them most useful are important and timely contributions as policy makers debate new IoT privacy and security regulations.

Our contributions are as follows:

1. An understanding of IoT device purchasers’ conceptions, misconceptions, and concerns about device privacy and security and the steps they take to address their concerns.
2. Identification of latent, unprompted privacy and security concerns, and distinctions between active behaviors toward privacy concerns and passive attitudes toward security risks.
3. A prototype IoT device privacy and security label, qualitative observations on its use, and recommendations for effective label design.

2 BACKGROUND AND RELATED WORK

In this section, we outline prior research on consumer purchase behavior and how it is impacted by privacy and security. We also present related work on labels.

**Consumer Purchase Behavior**

Purchase behavior is defined as the set of decisions people make and the actions they take when buying and using a product [6]. Purchasing comprises seven stages: need recognition, information search, pre-purchase alternatives evaluation, purchase, consumption, post-consumption evaluation, and divestment [8]. Pre-purchase behavior involves deciding on what to buy and when to buy a product [68], whereas post-purchase behavior includes steps consumers take to compare their expectation of the product to their perceived reality and manage their concerns and dissatisfaction [3, 34].

Researchers have identified factors that impact consumer choice in the pre-purchase evaluation stage. For instance, price, brand, features, aesthetics, and usability influence mobile phone purchases [41, 50, 51, 66]. The perceived quality of a product has been identified as the main driver of consumers’ purchase intentions [60]. Digital and social media have also been shown to impact consumers’ purchase behaviors [64]. In addition, word of mouth and reviews have been identified as influential factors [40, 65, 80]. In our interview study, we observed the impact of some of the previously mentioned factors on people’s IoT device purchasing decisions.

Studies have found that people are concerned about the privacy of their personal data when making online purchases [14, 38, 72]. Tsai et al. found that when accessible privacy information is made available in search results, consumers are more likely to purchase from privacy-protective websites, even if they are more expensive [71]. Kelly et al. found that adding concise privacy information to a mobile app store can impact users’ app-selection decisions [44].

**Labels**

Labels have been widely used to present information to consumers prior to making a purchase. For instance, consumers in the United States can compare the nutritional value of food products based on their nutrition facts label or they can compare the energy efficiency of light bulbs by looking at the Lighting Facts label. In addition, researchers have developed privacy labels for websites and have shown that people prefer privacy labels to traditional privacy policies [43].

Many consumers are concerned about the privacy and security of their IoT devices and want more transparency about how companies are collecting and using their data [37]. Moreover, experts warn about IoT device security vulnerabilities [5, 35, 69] that could allow an attacker to control a device or collect private data [15, 17, 18, 61, 79]. These vulnerabilities include insecure authentication mechanisms [62], transmitting unencrypted data [7, 78], and failure to promptly patch known bugs [36]. In addition, some devices collect sensitive information and transmit it to the device manufacturer or other parties, raising privacy concerns [4, 42, 53].
It is currently difficult for consumers to obtain information about the security and privacy of devices prior to purchase or at the time of purchase. The Mozilla “Privacy Not Included” buyers guide website is an example of a resource for consumers to look up privacy and security information for IoT devices [56]. However, it is not designed as a label and is not attached or linked to devices in a store. In addition, manufacturers do not always disclose all of the information included for each device. Moreover, as far as we know, the buyers guide has not undergone user testing.

Recent efforts are promoting the development and use of labels for IoT devices. For example, a UK government report recommended a voluntary labeling scheme for IoT devices [19] and a consortium of British universities is developing a consumer security index for IoT devices [63]. In the U.S., a number of lawmakers have introduced bills related to IoT device privacy and security. For example, the Cyber Shield Act of 2017 [49, 54] would create a voluntary label for IoT devices with independent testing and cybersecurity compliance grades. The Internet of Things Cybersecurity Improvement Act of 2017 [75] would direct government agencies to include contractual clauses that require security features for IoT devices purchased by the US government.

3 METHODOLOGY

We conducted a 24-participant semi-structured interview study followed by a 200-participant MTurk survey. The complete lists of interview and survey questions are provided in the online Appendix.

Semi-Structured Interview Study

We conducted semi-structured interviews in our lab with one or two interviewers present. We audio recorded all interviews and had them transcribed by a transcription service.

Recruitment, selection, and compensation. To recruit participants for our one-hour interview study, we posted flyers and advertised on Reddit, Craigslist, our institution’s study recruitment website, and social media. Participants were required to be at least 18 years old, and to have purchased at least one IoT device (smart home device or wearable) themselves. Our screening survey asked participants some demographic questions and questions about their IoT device(s), such as what devices they have and how they acquired them. We used the screening survey to exclude people who did not meet our criteria and to select a diverse sample (based on age, gender, occupation, and technical background). We invited selected participants to our lab for an interview and compensated each participant with a $25 Amazon gift card.

Pre-purchase behavior. We asked interviewees to tell us what IoT devices they have purchased, how long they have owned them, and why they purchased them. We asked them whether they had ever considered buying an IoT device and ended up not buying it, and the reasons for that decision.

We then asked interviewees about each IoT device they had purchased. We asked whether they purchased the device online or in a store, and the factors they considered before making the purchase. We wrote down each mentioned factor on a separate card to use later in the interview.

Post-purchase behavior. We asked interviewees about their post-purchase concerns and how they managed them.

IoT device privacy and security. The interviewer did not mention privacy or security until after the discussion of pre-purchase and post-purchase behaviors in order to avoid biasing interviewees. We asked interviewees to define privacy as it relates to IoT devices. We then asked whether they had any privacy concerns related to their devices and how they managed those concerns. Next, we asked them to define security and discuss any security-related concerns.

Value of privacy and security in purchase decisions. We asked interviewees to explain how important it is for them to know about the privacy and security of IoT device(s) they are considering for purchase. To further investigate consumers’ perceived value of privacy and security, we asked them to specify how much more they would be willing to pay to purchase an IoT device that provided privacy and security information as compared to one that did not.

We asked interviewees how comfortable they are with the data collected by their IoT devices. We also asked them to report whether they had ever read a privacy policy for their devices, how much they know about the privacy and security of their devices, and what they most want to know.

Finally, we presented our interviewees with a set of cards, each with one of the factors mentioned during the interview. We included cards for brand, price, privacy, and security, even if the interviewee had not mentioned these factors. We asked interviewees to sort the cards according to how much influence each factor had on their purchase decision.

Privacy and security label evaluation. Important privacy and security factors related to IoT devices have been identified previously [21, 25, 47, 48, 70]. We designed rough paper prototype labels based on familiar food nutrition labels to present these previously-identified privacy and security factors for three hypothetical IoT devices: a security camera, a smart toothbrush, and a smart thermostat. For each smart device, we designed three different variations of the label. In one label, we tuned the privacy and security information in such a way to make participants more comfortable with the data collection. For instance, we set the retention time to be as soon as the account on the device is deleted. In another label, we modified the values of the factors so that participants would feel less comfortable with the data collection.
(retention was forever, granularity was identifiable). For the third label, we tried to convey a trade-off. Figure 1 shows a label for a hypothetical security camera with poor privacy and security practices.

Eight participants saw three variations of labels for a security camera, eight participants saw three variations for a toothbrush, and eight participants saw three variations for a smart thermostat. We asked interviewees to think aloud as they compared the labels. We then asked them which device they would buy and what information on the labels helped them to make that decision.

We probed interviewees’ understanding of the information on the labels by asking them to go through one of the labels line by line and tell us what they believe it conveyed. We asked them to circle the parts which they found confusing. We then asked them which factors they consider most important, which information could be removed from the label, and whether there was any information they would like to see added.

At the end of the interview, we asked interviewees whether a privacy and security label would likely influence their IoT device purchase decisions. We asked about how they would want to be presented with the label while shopping online or in a store. We also asked about the importance of knowing about publicly-reported security vulnerabilities prior to purchasing IoT devices.

**Follow-Up Survey**

To be able to measure the reported influence of security and privacy on IoT device purchase decisions, we ran a supplementary MTurk survey with 200 participants from the United States. In this survey, we asked participants to imagine themselves engaging in three hypothetical comparison shopping scenarios for a security camera, a smart thermostat, and a smart toothbrush. We then presented our participants with 16 factors we found to be important from the interview study and asked them to rate each factor on a 5-point scale from no influence at all to a lot of influence. In addition, we asked participants whether they had purchased any IoT devices at all, as well as whether they had purchased any of the three types of devices we asked them about. At the end of the survey, we asked demographic questions.

**Data Analysis**

One of the researchers was the primary coder, responsible for creating and updating the codebook. To analyze the interview data, we applied structural coding to the interview transcripts. Structural coding, which is particularly useful for semi-structured interview studies, applies content-based phrases to a segment of data [52, 67]. We came up with eight structural codes (e.g., reasons to purchase smart home devices, privacy definition), which we divided into 61 subcodes.

The codebook was reviewed and revised by the researchers and then each interview was independently coded by two researchers. The final structural codes and subcodes can be found in the online Appendix. After resolving the coding disagreements, we achieved an inter-coder agreement of 91% Cohen’s Kappa. Kappa over 75% is regarded as excellent agreement [24]. For the remaining disagreement, we report the results of the primary coder.

Since the interview study is qualitative in nature and our sample size is small, we refrain from reporting the exact number of participants when presenting most of the results in Section 4. However, to provide readers with some sense of frequency, we adopt a consistent terminology, illustrated in Figure 2, to report these numbers.

Our MTurk survey was a repeated measure within-subject study. The dependent variables (DV) in our analysis were the scores from the 5-point scale, ranging from no influence at all to a lot of influence. We treated our DVs as interval
scales [57]. We evaluated the influence of security and privacy for each of the three devices; we did not evaluate the other 14 factors, which we had asked about so participants would not know that our interest was in security and privacy. The independent variables (IV) were demographic information, and information related to the IoT devices we tested in our survey. We also checked for the significance of the two-way interaction terms. To analyze the extent of the privacy and security influence on IoT device purchase decisions, we applied linear mixed-effect regressions with random intercept for each user to count for within-user data dependencies. The goal of the regression was to determine which independent variables would significantly associate with a change in influence. We used a significance threshold of 0.05.

**Limitations**

Participants in interview studies are prone to potential biases [1]. In our semi-structured interview, we avoided asking any leading question, mentioning security or privacy in the early parts of the interview, or correcting incorrect definitions or misconceptions. Even so, participants may have expressed more concern towards privacy and security later in the interview as they inferred the focus of our study.

It is important to note that our study focuses on consumer purchase behavior, and that our results may be less applicable to business purchase decisions and labels designed for corporate decision makers. With respect to purchase decisions, consumers and businesses are different in many ways [2, 45, 76]. Consumers make purchase decisions for personal consumption, whereas organizations make purchase decisions for the benefit of the business, and are more likely to purchase devices in bulk. Furthermore, corporate decision makers may consult a security expert in their organization before making IoT device purchase decisions.

**Ethics**

This study was reviewed and approved by our institution’s Institutional Review Board. All participants provided their informed consent to participate in the study, have their voice recorded, and have the recordings transcribed by a third-party company. We stored all digital files on a password-protected server and all paper files in a locked cabinet. The transcription company used a secure protocol to transfer files.

**4 RESULTS**

We present the results of our interview study here, following the flow of the interview. We discuss interviewees’ IoT device purchase behaviors and the evaluation of our privacy and security label. Finally, we report our follow-up survey results.

**Interviewees and Their Devices**

A total of 115 participants completed the online screening survey for our interview study. Of those, 99 participants were qualified and we invited a diverse sample of 25 participants to our lab for an interview. We excluded data from one of the interviewees, who revealed during the interview that she did not fully satisfy our study requirements. We analyzed the data from the remaining 24 participants. Our interviewees consisted of 14 female and 10 male participants with an average age of 36 years (std. dev. = 12 years). Eight interviewees had technical backgrounds. Our interviewees had a broad range of IoT devices. Information about our interviewees and their IoT devices is presented in Table 1.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Gender</th>
<th>Age</th>
<th>Tech Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>F</td>
<td>25-34</td>
<td>Y Camera</td>
</tr>
<tr>
<td>P2</td>
<td>F</td>
<td>35-44</td>
<td>Y Camera, doorbell, lights, smartwatch, IPA, TV</td>
</tr>
<tr>
<td>P3</td>
<td>M</td>
<td>25-34</td>
<td>Y IPA</td>
</tr>
<tr>
<td>P4</td>
<td>M</td>
<td>18-24</td>
<td>Y Smartwatch</td>
</tr>
<tr>
<td>P5</td>
<td>M</td>
<td>35-44</td>
<td>Y IPA, smartwatch, doorbell, lock, switches</td>
</tr>
<tr>
<td>P6</td>
<td>M</td>
<td>25-34</td>
<td>N Lights, activity tracker</td>
</tr>
<tr>
<td>P7</td>
<td>F</td>
<td>55+</td>
<td>N Lights, activity tracker, scale</td>
</tr>
<tr>
<td>P8</td>
<td>F</td>
<td>25-34</td>
<td>N IPA, switches, lights, smartwatch</td>
</tr>
<tr>
<td>P9</td>
<td>F</td>
<td>25-34</td>
<td>N Camera, TV</td>
</tr>
<tr>
<td>P10</td>
<td>M</td>
<td>18-24</td>
<td>Y IPA, activity tracker</td>
</tr>
<tr>
<td>P11</td>
<td>F</td>
<td>18-24</td>
<td>N IPA</td>
</tr>
<tr>
<td>P12</td>
<td>F</td>
<td>45-55</td>
<td>N IPA, activity tracker</td>
</tr>
<tr>
<td>P13</td>
<td>F</td>
<td>45-55</td>
<td>N IPA</td>
</tr>
<tr>
<td>P14</td>
<td>M</td>
<td>55+</td>
<td>N IPA</td>
</tr>
<tr>
<td>P15</td>
<td>M</td>
<td>25-34</td>
<td>Y Smartwatch</td>
</tr>
<tr>
<td>P16</td>
<td>F</td>
<td>25-34</td>
<td>N Activity tracker, TV</td>
</tr>
<tr>
<td>P17</td>
<td>F</td>
<td>55+</td>
<td>N TV, IPA, switches, activity tracker</td>
</tr>
<tr>
<td>P18</td>
<td>F</td>
<td>25-34</td>
<td>N IPA</td>
</tr>
<tr>
<td>P19</td>
<td>M</td>
<td>25-34</td>
<td>N Camera, lights, TV, thermostat, smoke alarm, activity tracker</td>
</tr>
<tr>
<td>P20</td>
<td>F</td>
<td>25-34</td>
<td>N Activity tracker</td>
</tr>
<tr>
<td>P21</td>
<td>F</td>
<td>25-34</td>
<td>N Smartwatch, IPA</td>
</tr>
<tr>
<td>P22</td>
<td>M</td>
<td>45-55</td>
<td>N Smartwatch, IPA</td>
</tr>
<tr>
<td>P23</td>
<td>M</td>
<td>35-44</td>
<td>Y Smoke alarm, IPA, camera, thermostat, switches, activity tracker</td>
</tr>
<tr>
<td>P24</td>
<td>F</td>
<td>35-44</td>
<td>N IPA, TV</td>
</tr>
</tbody>
</table>

Table 1: Participant demographics and IoT devices. IPA stands for Intelligent Personal Assistant (e.g., Amazon Echo, Google Home, Apple HomePod).
We asked interviewees to define privacy and security specifically about IoT devices. Their definitions demonstrated that they had a narrow and limited knowledge of privacy and security, and some could not distinguish between them.

Most interviewees defined privacy related to smart devices as having control over personal data. For example, P16 said:

“privacy is whether it’s up to me or them how they use my data.” Some mentioned who data is being shared with and a few talked about types of data being collected, retention time, purpose of data collection, and inferred data.

When we asked interviewees to define security related to IoT devices, most of them mentioned protection from unauthorized access (“being hacked”). Half of the interviewees talked about means of protection. For instance, some mentioned password protection and authentication and a few talked about firewalls, encryption, and physical locks. A few mentioned risks associated with unauthorized access to personal data.

In general, our interviewees felt more power toward privacy as opposed to security. When defining privacy, participants mentioned that they should have control over their data, whereas they were mostly passive when defining security as the device getting hacked. About half of our interviewees were not able to differentiate between privacy and security of smart devices. However, most of the interviewees who mentioned having pre-purchase or post-purchase privacy or security concerns were more able to differentiate correctly. This suggests that a lack of privacy or security concerns might be attributed to not having correct and distinct definitions for these two concepts.

### Purchase Behavior Categories

<table>
<thead>
<tr>
<th>Purchase Behavior</th>
<th>Awareness</th>
<th>Knowledge</th>
<th>Evaluation</th>
<th>Concern</th>
<th>Management</th>
<th>Participant ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wise Proactive Protective</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>P2, P3, P5, P8, P10, P20, P23</td>
</tr>
<tr>
<td>Cautious Proactive Protective</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>P15, P19</td>
</tr>
<tr>
<td>Wise Passive Protective</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>P1, P10, P24</td>
</tr>
<tr>
<td>Cautious Passive Protective</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>P11, P13, P17</td>
</tr>
<tr>
<td>Wise Passive</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>P14, P16, P18, P22</td>
</tr>
<tr>
<td>Cautious Passive</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>P12</td>
</tr>
<tr>
<td>Unconcerned</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>✔️ ✔️ ✔️ ✔️</td>
<td>P4, P6, P7, P9, P21</td>
</tr>
</tbody>
</table>

Table 2: Seven purchase behavior categories and the participants whose behaviors are described by each.

### Pre-Purchase Behavior

Curiosity was a primary reason for purchasing IoT devices, especially for owners of Intelligent Personal Assistants (IPAs) such as Amazon Echo and Google Home. Wearable purchasers were primarily motivated by a desire to improve health and fitness. Price and convenience were other reasons mentioned frequently by interviewees.

Most interviewees mentioned reliability concerns and lack of necessity as reasons not to buy smart home devices, and price as a reason not to buy wearables. Some people mentioned privacy and security concerns as reasons they avoided purchasing a specific smart home device. However, only a few specified privacy concerns as a reason not to buy wearables. P6 mentioned that he did not buy a smart door lock because he was not comfortable with the security of the device. Moreover, P7 said she would not buy Google Home due to concerns that it would listen to her all the time.

Participants mentioned 16 factors that influenced their purchase decisions: look and feel, customer service, prior experience with the device or similar devices, ease of use, reliability, opinion from experts (magazine reviews, electronics store employees), compatibility with other devices, durability, opinion from friends, opinion from family members, brand, privacy, security, customer reviews, price, and features. From the card sorting activity, we found that interviewees ranked privacy and security as the most influential factors after price and features. The card sorting activity should not be interpreted quantitatively due to both the small number of participants and the difference in the number of cards sorted by each participant. To further explore the relative influence of factors, we conducted a large-scale survey.

### Post-Purchase Behavior

When we asked interviewees about any concerns and issues they had with their IoT devices, most reported minor technical issues and about half mentioned privacy or security concerns. Note that at this point in the interview the interviewer had not yet mentioned privacy or security.

Interviewees who reported privacy concerns were almost all concerned about IPAs or smart TVs listening to them. Those who reported security concerns described mitigation steps they took, such as connecting their IoT devices to a router separate from the rest of their home network.

### Defining IoT Device Privacy and Security

We asked interviewees to define privacy and security specifically about IoT devices. Their definitions demonstrated that they had a narrow and limited knowledge of privacy and security, and some could not distinguish between them.

Most interviewees defined privacy related to smart devices as having control over personal data. For example, P16 said:
privacy and security of their devices after purchase, and took actions to manage their concerns (e.g., by updating the system frequently, using a password generator, changing the position of a home camera, using a separate router for IoT devices, and turning off/muting the device). We labeled this behavior as proactive protective. Most people who exhibited proactive protective behavior were aware of risks and knowledgeable about privacy and security (labeled as wise). However, some were aware of risks but provided incorrect or indistinct definitions of security and privacy (labeled as cautious).

Most interviewees did not take privacy and security into account while making the purchase, but were concerned about their device privacy or security after the purchase. We found that post-purchase concerns were mostly caused by hearing about concerns from friends, media reports, and the device functioning in an unexpected way. We labeled this behavior as passive. Half of the interviewees who exhibited passive behavior took some actions to manage their concerns (labeled as passive protective). P1 reported that she managed concerns about her laptop camera but was unable to manage concerns about her home security system: "so in a movie, you know, some crazy hacker, they can hack into all the films and cameras, so I know I put a sticker on my laptop camera, always, but I can’t put a sticker on my home camera because I need to see what’s happening, so I do worry about the system being hacked, my camera system being hacked."

Finally, a few of our interviewees reported being unconcerned about the privacy and security of their devices in both the pre-purchase and post-purchase stages. We noted two common reasons as to why people were not concerned about the specific IoT devices they had. They either did not perceive the collected data to be sensitive or expressed self-efficacy toward protecting themselves against the privacy or security related threats. For instance, P9 said she was unconcerned about her home camera system because “you can’t view it online or even on the app without the phone being connected to the camera and without having a user name and password.” Another unconcerned interviewee, P21, said “I have a passcode on it. So, I’m not worried about someone looking at it and besides my texts there’s not really anything that I feel needs to be private.” It is important to mention that being unconcerned does not necessarily imply having no privacy or security concern about any IoT devices, as some of the unconcerned interviewees said they would be concerned if they owned other types of IoT devices.

Value of Privacy and Security in Purchase Decisions

While only eight interviewees considered privacy or security as a factor in their comparison shopping (proactive protective), almost all said they would like to know about the privacy and security of devices before making future device purchases. Some noted that the importance of this information would depend on the type of data being collected by the IoT device. Interviewees were most interested in knowing about the purpose of data collection and privacy choices. We asked interviewees to specify what premium they would be willing to pay, if any, for a device with privacy and security information provided. Almost all interviewees said they would be willing to pay a premium of 10%-30% of the base price of the device. Reasons for their willingness to pay a premium included assurance that security and privacy would be protected and peace of mind. Those who were reluctant to pay more for privacy and security information often mentioned lack of trust in the device company providing the information. For instance, P12 said: "I wouldn’t necessarily believe it because, like with the Facebook thing, regardless of what they say, they’re gonna have all that information." Among different purchase behavior categories, “proactive protectors” were willing to pay slightly more, as they were more concerned about their devices even prior to purchase. Other researchers have also shown that consumers are willing to pay a premium for privacy [20, 71]. However, it is not clear how much of a premium consumers are willing to pay. An incentive compatible study is needed to further elicit consumers’ willingness to pay.

Privacy and Security Label Evaluation

In the last section of the interview, we showed each participant labels for three hypothetical IoT devices and asked them to compare the devices and provide feedback on the labels. The focus of this evaluation was mostly on the contents of the labels, although we also received insightful suggestions on improving the design of the labels.

Almost all participants first compared the labels based on the ratings, and quickly identified the privacy protective label. On each label, there were ratings from an independent privacy lab, an independent IT security institute, and Consumer Reports (CR). Participants particularly liked having ratings from independent research labs. These ratings were especially of interest to those who previously reported their lack of trust in IoT companies. Nonetheless, participants wanted to know what factors went into the ratings. The CR score was regarded as an important piece of information mainly by those participants who were familiar with Consumer Reports and had previously consulted their reviews when making a purchase.

After participants compared the labels, we asked them about the privacy and security sections of the labels. Almost all participants reported that the labels covered all the topics they wanted to know about, and they especially liked the inclusion of information about choices. A few participants wanted to know where data storage servers were located.
Participants discussed their comfort level with the specific values shown for some of the fields on the labels. For example, almost all participants were comfortable with data being used for research, but some did not trust that companies would not also use their data for marketing. Almost all were uncomfortable when the retention time was forever, but were comfortable with companies retaining data until you delete your account. However, P5 recognized utility in longer data retention: “This is a thermostat, retention one month. You know what, actually, now that I think about it, the retention might be useful if you kept it forever, because you could do analytics across time.” Almost all participants preferred aggregated and anonymous data over identifiable data, although most participants could not distinguish between aggregated and anonymous information. P8 understood the difference and recognized the value of identifiable data: “So if the data being used is the aggregate behavior of me and all the people in my three-digit ZIP code, then that would be an empty feature for me if I want my thermostat to respond to when I’m home.”

Participants were more focused on the convenience of security factors than on their level of protection. For instance, almost all participants said they wanted “automatic updates” to be available because they found them more convenient than manual updates. In addition, almost all preferred fingerprint authentication over passwords due to convenience. Similarly, participants favored optional Internet connectivity over required connectivity because they wanted their devices to be able to function when Internet connectivity was unavailable.

We followed a user-centered design process and revised the label between interviews to address parts that were unclear to participants. For example, some participants did not understand the term “account information,” so we changed the term to “login info and device configuration.” In addition, we found that the term “granularity” confused participants, so we changed it to “level of detail.”

Participants found the final version of the labels to be understandable, easy to read, and useful. P11 compared the label to privacy policies: “As opposed to those long documents that you usually need to read, I think this is a very efficient way and I cannot think of a better way than this.” P24 pointed out the importance of being reminded of privacy and security at purchase time: “If you don’t know about the label, you don’t think, man, I just need to know the security and privacy things about this product before I buy it. You don’t think that.” Some participants noted that it had been difficult to find privacy and security information prior to purchasing an IoT device.

We discussed participants’ preferences for where to find the label when shopping online or in a store. Most participants wanted to have the label in the online store’s device description, as one of the images, or after the features and before the customer reviews. For in-store shopping, about half of the participants wanted the label to be on the shelf to compare devices easily, even though some participants noted the possibility of devices being placed incorrectly in the store.

We asked interviewees whether they would like to be presented with publicly reported vulnerabilities of smart devices before making a purchase. Almost all of the respondents reported that they would like to have this information in a label and that it might impact their purchase decision. Interviewees particularly wanted to know how serious those incidents were and how prompt the manufacturers were to fix security problems.

**Follow-Up Survey**

For our supplementary within-subject survey, we recruited 200 MTurk participants from the United States. 87 participants reported that they had purchased at least one IoT device themselves, of which 28 reported purchasing a smart security camera, 28 a smart thermostat, and 22 a smart toothbrush. 118 participants reported being female, 81 reported being male, and one reported other. The average age of our participants was 38 years (std. dev. = 10 years). 44 participants reporting having a technical background.

We found that the importance of privacy and security depended on the type of device. As shown in Figure 3, most participants said security and privacy would influence their purchase of security cameras and smart thermostats, but not smart toothbrushes. Our regression results indicated that the influence of security and privacy information was significantly higher for a security camera ($p < 0.05$) than for a smart thermostat. For the toothbrush, privacy and security information were significantly less influential ($p < 0.05$) compared to the other two devices. The summary statistics
While about half the interviewees brought up privacy or security concerns and help consumers consider privacy and security in their purchase decisions and device use. We then discuss several design considerations for more effective privacy and security labels. Finally, we discuss approaches to label adoption and mechanisms to promote label accuracy.

### Latent Concern

While about half the interviewees brought up privacy or security concerns before we mentioned them, the other half did not discuss privacy or security until our prompt. However, once prompted, almost all interviewees reported being concerned about the privacy and security of their IoT devices. This suggests that for some consumers, privacy and security are latent concerns, which can be surfaced readily if privacy and security information is made salient, for example by appearing in a label. Once consumers are prompted to consider privacy and security information, they may incorporate it into their purchase decision process.

Designers of privacy and security labels should more effectively communicate risks to consumers. One approach to better convey the relative risks of privacy and security is to combine data types with their purposes. In our study, some participants had difficulty relating data with their purpose. Another design idea is to distinguish expected and unexpected data practices. Expected practices are the data collections which are necessary for the core functionality of the device, whereas unexpected practices include non-essential data collection or use, such as selling data to third parties or profiling users for targeted advertising.

Privacy and security information can also shape consumer behavior after a device is purchased. Labels may inform consumers about their privacy and security choices and how to manage them. They may also make them aware of potential privacy or security vulnerabilities that they may be able to mitigate themselves by engaging in protective measures (e.g., turning off a device when not in use or positioning a device so as to avoid collecting data in a private space). Our prototype label indicated when privacy or security choices were available to consumers. In addition, we provided them with protips to suggest protective privacy and security behaviors. To further help consumers make informed privacy and security decisions based on the protips, designers can provide consumers with an understandable user manual on how to implement them.

### Label Design Considerations

Some interviewees requested more information on our prototype label to make an informed purchase decision, such as definitions of some of the terms, encryption protocols used, and information about the process the independent privacy and security labs followed to rate the IoT devices. Having more information could be particularly useful for "cautious" consumers, as they have little knowledge of privacy and security in the context of IoT devices. While adding all of this information to a static label would likely reduce its usability, additional information can be included in an interactive online label, in which consumers can hover over or click on each factor to obtain additional information. The QR code on a printed static label can direct consumers to an online interactive version. This "layered" approach has been recommended for privacy notices [59, 77]. However, it is important that the static version of the label (the top layer) contain the most critical information, as it is likely that most consumers will glance over labels without interacting with them.

When comparing IoT devices, privacy and security star ratings immediately caught the attention of almost all interviewees. Aside from being a glanceable synopsis of key privacy and security factors, ratings were attractive due to the independence of the organizations (e.g., Consumer Reports) that provided the ratings. They were especially favored by interviewees who mistrusted the manufacturers and questioned whether they would adhere to their claims. Security ratings may help mitigate consumers' common misunderstandings around security information.

Throughout our interviews, we observed that participants discussed their active control over privacy, but seemed resigned to not being able to control security. While users may feel empowered to take physical steps to protect privacy (e.g., by covering a camera lens), they may view security as an innate, uncontrollable property of the device, or they may lack knowledge to understand the actual security risks or how to mitigate them. Such passive attitudes toward security factors were common across purchase behavior categories.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Coeff.</th>
<th>Std. Err.</th>
<th>t-value</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (DV: security)</td>
<td>4.03</td>
<td>0.08</td>
<td>46.41</td>
<td>4.03</td>
<td>0.97</td>
</tr>
<tr>
<td>Smart thermostat</td>
<td>-0.61</td>
<td>0.09</td>
<td>-6.41</td>
<td>3.47</td>
<td>1.29</td>
</tr>
<tr>
<td>Smart toothbrush</td>
<td>-1.56</td>
<td>0.09</td>
<td>-16.54</td>
<td>2.52</td>
<td>1.41</td>
</tr>
<tr>
<td>Intercept (DV: privacy)</td>
<td>3.91</td>
<td>0.09</td>
<td>42.66</td>
<td>3.91</td>
<td>1.11</td>
</tr>
<tr>
<td>Smart thermostat</td>
<td>-0.53</td>
<td>0.09</td>
<td>-5.80</td>
<td>3.37</td>
<td>1.29</td>
</tr>
<tr>
<td>Smart toothbrush</td>
<td>-1.35</td>
<td>0.09</td>
<td>-14.69</td>
<td>2.55</td>
<td>1.45</td>
</tr>
</tbody>
</table>

Table 3: Summary statistics and regression results of the reported influence security and privacy have on participants’ purchase decisions. There were 600 observations for each regression (200 responses for each smart device) and the baseline for both regressions is the smart security camera. The factors in the table are all statistically significant ($p < 0.05$).
Even some “wise” participants viewed security mitigation as overly burdensome. Thus, we found that interviewees were using the information in the security section of our labels to make security decisions that were more about convenience than security. Our results suggest that the design of the security portion of the label should bring out security risks and their implications more directly (e.g., communicate that when data is transmitted without encryption, it may be accessible to eavesdroppers). Adopting more robust security practices may not always be convenient for consumers, even if well explained. Thus it is important for IoT device manufacturers to find ways to provide security without burdening users, and to make more secure options the default.

Label Adoption and Enforcement

In order for labels to be practically useful, they need to be widely used and convey accurate information. Use of labels may be mandated by regulations or strongly encouraged through “safe harbor” provisions. Even in the absence of regulatory mandates, retailers may require labels on products they sell or may promote products that have labels. Some manufacturers may adopt labels voluntarily to gain consumer trust. As P20 stated: “I would definitely trust something that had this above something that didn’t.”

Some interviewees reported trusting well-known brands of IoT devices more than unfamiliar brands. This is consistent with prior work showing the impact of company size and reputation on consumer trust [39, 74] and purchase behavior [16, 55]. As a result, smaller and less well-known companies will likely take longer to develop consumer trust. However, a label may help level the playing field by allowing companies to be transparent about the privacy and security of their devices, and display independent ratings that may reassure consumers.

While we have described several approaches to mandating or encouraging label adoption, it should be noted that past efforts to encourage standardized privacy disclosures have faltered in the absence of regulatory mandates [13].

Enforcement mechanisms are needed to ensure that there are consequences for companies that convey inaccurate information on their labels. In the United States, the Federal Trade Commission or state attorneys general would likely prosecute companies who are found to make false claims on their labels, similar to what happens when companies are found to make false claims in their privacy policies [28–32]. In Europe and other countries around the world, enforcement actions could likely be taken by data protection commissioners.

The rapid pace at which IoT devices receive software and firmware updates could make it a challenge for manufacturers to keep their labels up to date. This also means that the adherence of the IoT devices’ actual behavior to what is on the label is a moving target as features are added or removed, bugs introduced or fixed, and the firmware updated. A realistic solution is to have labels marked with software and hardware version numbers, with QR codes or hyperlinks to the label for the latest firmware.

Finally, it is important to recognize that some security and privacy issues may be best addressed by mandating or prohibiting certain practices, rather than simply disclosing practices on a label and leaving it to consumers to avoid IoT devices with egregious security or privacy flaws.

6 CONCLUDING REMARKS

We conducted an in-depth semi-structured interview study with participants who have purchased at least one IoT device, to explore their knowledge and behavior regarding IoT security and privacy. Some participants considered privacy and security while making IoT device purchase decisions, while half of them were concerned about device security and privacy only after the purchase. Almost all participants acknowledged the importance of having privacy and security information, and said they would pay a premium to have this information available at purchase time. Most participants in a followup MTurk study said security and privacy were factors that would influence their purchase decisions for some types of IoT devices, especially those they perceive as collecting sensitive information. Finally, we developed a prototype IoT device privacy and security label. Our interviewees found the design to be understandable, and especially appreciated the inclusion of independent ratings and privacy choices. However, most did not use the detailed security information to minimize security risk. We discuss design considerations for IoT security and privacy labels and paths to adoption and enforcement.

ACKNOWLEDGMENTS

This research has been supported in part by DARPA and the Air Force Research Laboratory under agreement number FA8750-15-2-0277. The US Government is authorized to reproduce and distribute reprints for Governmental purposes not withstanding any copyright notation thereon. Additional support has also been provided by Google. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, the Air Force Research Laboratory, Google, or the US Government.

REFERENCES
