

Informing Privacy and Security Decision Making in an IoT World

Pardis Emami-Naeini

CMU-ISR-20-106

May 2020

School of Computer Science
Institute for Software Research
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:

Lorrie Faith Cranor (Co-Chair)

Yuvraj Agarwal (Co-Chair)

Lujo Bauer

Mohammad Reza Haghighat (Intel Corporation)

*Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Societal Computing.*

© 2020 Pardis Emami-Naeini

The research reported in this thesis has been supported in part by DARPA and the Air Force Research Laboratory FA8750-15-2-0277, and NSF awards TWC-1564009 and SaTC-1801472. Additional support has also been provided by Google and by the Carnegie Mellon CyLab Security and Privacy Institute. The views and conclusions contained in this document are those of the author, and should not be interpreted as representing the official policies, either expressed or implied, of any sponsoring institution, the U.S. government, or any other entity.

Keywords: Privacy, Security, Usability, Internet of Things (IoT), Decision making, Label

To those who never give up.

Abstract

In recent years, a massive number of devices have emerged with the capability to connect to the Internet, thereby providing people with unprecedented benefits. These Internet of Things (IoT) devices are increasingly used to improve energy efficiency, home security and convenience, and by 2025, it is estimated to have an installed base of 75 billion IoT devices throughout the world. The cybersecurity threats of these devices, however, are not as appealing as their benefits. Baby monitors get hacked, Amazon Echo devices send private conversations to others, and Samsung Smart TVs start recording without users' knowledge. One explanation for these overwhelmingly challenging risks of IoT devices could be overlooking privacy and security early on in the product life cycle due to lack of resources (e.g., expertise, money). Integrating privacy and security safeguards into IoT devices could reduce their risks or mitigate their potential harms. At the same time, IoT manufacturers are not transparent about their privacy and security practices, leaving consumers with little information when purchasing IoT devices. This lack of information at the time of purchase could result in people bringing home a vulnerable device and easily scaling up the threat by connecting the device to their home network.

Thanks to privacy and security experts and media reports, people are becoming aware of the threats of smart devices. However, despite growing concerns about the privacy and security of IoT devices, people have difficulty specifying their privacy and security preferences and considering them when making IoT-related purchase decisions. To enable informed decision making during the purchase process of IoT devices, we need to understand how people feel about the privacy and security implications of these devices. Moreover, effective ways of communicating important privacy and security factors to consumers of IoT devices need to be carefully studied.

In this thesis, we first explore the factors influencing users' privacy concerns and preferences toward data collection of smart devices. To this end, we quantify users' privacy preferences and expectations with the aim of statistically modeling privacy-related attitudes and reported behaviors by factors such as the collected data, the purpose of data collection, and the retention time. In a 1,007-participant online study, we found that participants are significantly more comfortable when seemingly innocuous information such as the room's temperature or their presence is being collected, as compared to when more sensitive information like their biometrics (e.g., fingerprints) are being collected. In addition, participants are significantly more

willing to allow data collection in a public space (e.g., library) than a private location (e.g., at home).

Next, we explore how users' IoT-related privacy decision making would be influenced when receiving social cues from privacy experts and friends. We found that both friends and privacy experts significantly impact participants' privacy-related decision making. Following our overarching goal to inform privacy-related decision making, we delve into designing a label to effectively inform consumers about the privacy and security practices of smart devices at the time of purchase. To achieve this, we first interviewed 24 IoT consumers on the factors they consider when purchasing smart devices and found that currently, seeking understandable privacy and security information for smart devices is difficult or impossible. This finding motivated us to seek an effective mechanism to inform consumers by better communicating this information at the point of sale. We proposed creating a usable privacy and security nutrition label for IoT practices, building on prior projects that have used nutrition labels in other privacy contexts. To explore the actual content of such a label, we conducted a study with experts from diverse domains and identified 47 privacy and security attributes to include on a two-layer label. Finally, we evaluated the efficacy of attribute-value pairs presented on the label in conveying risk to consumers as well as its effect on their willingness to purchase the smart device. Our results show that data privacy and security information is more powerful in swaying consumers' risk perception than changing their willingness to purchase.

Thesis statement: **The objective of this thesis is to establish a thorough understanding of how users make privacy-related decisions when interacting with IoT devices, combine the obtained knowledge with experts' insights to develop a privacy and security label for IoT devices, and finally evaluate its usability and risk communication to effectively inform consumers' IoT-related purchase decision making.**

Acknowledgments

My PhD journey has been full of ups and downs. The one thing that kept me going forward and never looking back was being surrounded by extremely kind, humble, and supportive mentors, who I did look up to and follow constantly.

Starting with my advisors Lorrie and Yuvraj: Lorrie Cranor is a true leader, who I learnt the meaning of grit from. Her work and lifestyle all showed me how to be productive in professional and personal aspects and at the same time always be energetic and open to new discussions. Yuvraj Agarwal taught me to be humble and always thrive to learn a new thing. His enormous optimism was all I needed when I felt down about the path forward. I never left his office feeling unhappy or perhaps I should say I never left his office not being filled with energy and confidence to be the best I can be. His words were life lessons.

I will always be thankful to my committee members Lujo Bauer and Mohammad Reza Haghighat for the lessons they taught me in my journey. Lujo, by his unique way of thinking, showed me how important it is to think and reason critically. He never stopped teaching me how details matter. Mohammad Reza is a mentor for life, who taught me to speak up and not to be shy. I will never forget him telling me, during my internship at Intel, not to be afraid of sitting in the front row of the auditorium, when the CEO is giving a speech. He is the kind of person, who despite being enormously knowledgeable in different topics, is overwhelmingly down to earth and that is what I try to follow every day.

I would like to thank all my co-authors and collaborators at CMU, including Nicholas Christin, Norman Sadeh, and Hanan Hibshi. I have learnt a great deal from each one of them.

I am forever grateful to all the staff in Institute for Software Research (ISR) and CyLab. They were among the most committed people I have ever seen and their commitment made my PhD journey the smoothest possible. I am especially thankful to Karen Lindenfelser, whose kindness was always a bright light in the years I have spent here.

I am very thankful to my amazing mentors at Intel, Richard Chow and Heather Patterson, and the great students I had the pleasure to work with: Janarth Dheenadhayan, Shreyas Nagare, Henry Dixon, and Soopawat Vitoorapakorn. I am especially thankful to my wonderful friends in the CUPS lab: Jessica Colnago, Josh Tan, Hana Habib, Maggie Oates, Aurelia Augusta, Sarah Pearman, Kyle Crichton, Abby

Marsh, and Martin Degeling.

I have had the privilege to learn from great statisticians at CMU, Alex Davis and Howard Seltman, whose guidance I used throughout my thesis.

I had the amazing opportunity to spend one of my summers during my PhD as an Intern at Microsoft Research and learn from some of the smartest researchers I have ever seen, which I am deeply grateful for. I am particularly thankful to my mentor at MSR, Kim Laine, whose vast knowledge and constant support made my internship an unbelievably valuable and rewarding experience. He was a natural teacher and mentor, who taught the most difficult concepts in the most comprehensible way.

I would like to acknowledge the unforgettable and continuous support I received from my undergraduate mentors in Physics and Computer Engineering Departments at Sharif University of Technology, Mohammad Akhavan Farshchi, Shahin Hessabi, and Ali Movaghar. Their genuine trust in me was the fuel to reach for the moon.

Doing a PhD at Pittsburgh and CMU would have not been the same experience without my closest friends, Sruti Bhagavatula, Mahmood Sharif, Janos Szurdi, Orsi Kovacs, Aymeric Fromherz, Steve Matsumoto, Amirbehshad Shahrabi, and Naji Shajari.

I am also especially thankful to my far away friends, Aida Arman Moghadam, Milad Asgari, and Vahid Gerayi Nezhad, who always supported me in this journey and washed away all the sad and lonely moments with their humor.

I am forever grateful for the incredible love, support, and patience of my family. There is no word to describe how much I appreciate what my parents did for me over the past 27 years of my life and will continue to be doing. My sister and brother, Parisa and Saeed, have always reminded me that I am much more powerful than any difficulty I face. I am truly grateful to my brother-in-law, Iman, who filled my CMU PhD application form while I was drowning in completing my application materials. It is no exaggeration to say that this thesis would have not been a reality without that application being submitted. For the past year, no single day has passed without my belly laughs at the delightful pictures and videos of my incredibly “googooli” niece, Nika.

Final thanks go to Navid Naderi. His never-ending support and encouragement all these years made me believe in myself to accept new challenges and never give up. His calm manner changed the way I looked at everything and helped me believe that no matter how hard the problems are, I can find not only one way, but many ways to solve them, without losing sleep over.

Chapter 3 is a lightly edited version of a paper previously published as: Pardis Emami-Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, Norman Sadeh. Privacy Expectations and Preferences in an IoT World. In Proceeding of the 13th Symposium on Usable Privacy and Security (SOUPS), 2017 [115]. The dissertation author is the primary investigator and author of this paper.

Chapter 4 is a lightly edited version of a paper previously published as: Pardis Emami-Naeini, Martin Degeling, Lujo Bauer, Richard Chow, Lorrie Faith Cranor, Mohammad Reza Haghghat, Heather Patterson. The Influence of Friends and

Experts on Privacy Decision Making in IoT Scenarios. In Proceeding of the 21st ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW), 2018 [116]. The dissertation author is the primary investigator and author of this paper.

Chapter 5 is a lightly edited version of a paper previously published as: Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, Lorrie Faith Cranor. Exploring How Privacy and Security Factor into IoT Device Purchase Behavior. In Proceeding of the 37th ACM Conference on Human Factors in Computing Systems (CHI), 2019 [117]. The dissertation author is the primary investigator and author of this paper.

Chapter 6 is a lightly edited version of a paper previously published as: Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, Hanan Hibshi. Ask the Experts: What Should Be on an IoT Privacy and Security Label?. In Proceeding of the 41st IEEE Symposium on Security and Privacy (S&P), 2020 [114]. The dissertation author is the primary investigator and author of this paper.

Chapter 7 is a lightly edited version of a paper currently under submission as: Pardis Emami-Naeini, Janarth Dheenadhayalan, Yuvraj Agarwal, Lorrie Faith Cranor. Which Privacy and Security Attributes Most Impact Consumers' Risk Perception and Willingness to Purchase IoT Devices?. The dissertation author is the primary investigator and author of this paper.

Contents

- 1 Introduction** **1**

- 2 Background and Related Work** **5**
 - 2.1 Privacy Decision Making 5
 - 2.1.1 Factors Impacting Privacy Preferences and Concerns 6
 - 2.1.2 Decision Making under Social Influence 7
 - 2.1.3 Predicting Privacy Preferences 9
 - 2.2 Purchase Decision Making 9
 - 2.2.1 Purchase Process and Willingness to Purchase 9
 - 2.2.2 Factors Impacting Purchase Decision 10
 - 2.3 Risk Perception 10
 - 2.4 Risk Communication 11
 - 2.4.1 Labels 11
 - 2.4.2 Privacy and Security Guidelines and Best Practices 12

- 3 Privacy Expectations and Preferences toward IoT Data Collections** **15**
 - 3.1 Methodology 16
 - 3.1.1 Factors Impacting Preferences 18
 - 3.1.2 Predicting Preferences 20
 - 3.1.3 Qualitative Analysis of Preferences 22
 - 3.1.4 Limitations 23
 - 3.2 Results 23
 - 3.2.1 Participants 23
 - 3.2.2 Comfort with Data Collection 23
 - 3.2.3 Allowing or Denying Data Collection 27
 - 3.2.4 Data Collection Notification Preferences 30
 - 3.3 Discussion 34
 - 3.3.1 Privacy Preferences Are Complex 34
 - 3.3.2 Addressing Privacy Concerns 34
 - 3.3.3 Towards Awareness and Control 35
 - 3.4 Conclusion 36

- 4 The Influence of Friends and Experts on Privacy Decision Making in IoT Scenarios** **37**
 - 4.1 Methodology 39
 - 4.1.1 Study Design 39
 - 4.1.2 Data Analysis 41
 - 4.1.3 Free Text Responses 43

4.1.4	Limitations	43
4.2	Results	45
4.2.1	Faster Privacy-Related Decision Making	45
4.2.2	Inferred Influence	47
4.2.3	Reported Influence	48
4.2.4	Willingness to Trust Influencers	49
4.3	Discussion	51
4.3.1	Privacy Experts or Friends?	52
4.3.2	Wisdom of Crowds	53
4.3.3	Social Influence in Action	53
4.4	Conclusion	55
5	Exploring How Privacy and Security Factor into IoT Device Purchase Behavior	57
5.1	Methodology	58
5.1.1	Semi-Structured Interview Study	59
5.1.2	Follow-Up Survey	60
5.1.3	Data Analysis	62
5.1.4	Limitations	62
5.2	Results	63
5.2.1	Interviewees and Their Devices	63
5.2.2	Pre-Purchase Behavior	63
5.2.3	Post-Purchase Behavior	64
5.2.4	Defining IoT Device Privacy and Security	65
5.2.5	Purchase Behavior Categories	65
5.2.6	Value of Privacy and Security in Purchase Decisions	66
5.2.7	Privacy and Security Label Evaluation	67
5.2.8	Follow-Up Survey	68
5.3	Discussion	69
5.3.1	Latent Concern	69
5.3.2	Label Design Considerations	70
5.4	Conclusion	71
6	Ask the Experts: What Should Be on an IoT Privacy and Security Label?	73
6.1	Methodology	74
6.1.1	Expert Elicitation Study	75
6.1.2	Semi-Structured Interviews with Non-Expert Consumers	78
6.1.3	Limitations	80
6.2	Results	81
6.2.1	Definition, Assessment, and Accountability	82
6.2.2	Factors to Include in the IoT Label	84
6.2.3	Attitudes toward Labels and Layered Design	89
6.2.4	Prototype Privacy and Security Label	90
6.3	Discussion	90
6.3.1	Star Ratings vs. Certification Levels	93

6.3.2	Privacy and Security Evaluation and Scoring	93
6.4	Conclusion	96
7	Which Privacy and Security Attributes Most Impact Consumers' Risk Perception and Willingness to Purchase IoT Devices?	97
7.1	Methodology	98
7.1.1	Study Design	98
7.1.2	Data Analysis	102
7.1.3	Limitations	102
7.2	Results	103
7.2.1	Concern Level and Purchase History	103
7.2.2	Models to Describe Risk Perception and Willingness to Purchase	104
7.2.3	Qualitative Results	113
7.2.4	Decision Criteria to Assess Risk	117
7.2.5	Impact of the Extremes	117
7.3	Discussion	117
7.3.1	Label Content	117
7.3.2	Information Presentation	119
7.3.3	Viewing Label as a Whole	120
7.4	Conclusion	120
8	Conclusion and Future Work	123
8.1	Summary of the Discussed Research	123
8.2	On the Usefulness of Labels	124
8.3	International Labeling Efforts	125
8.3.1	United Kingdom	125
8.3.2	Finland	125
8.3.3	Singapore	126
8.3.4	Other International Activities	127
8.4	Specification and Tool Accompanying the Label	128
8.4.1	Specification for Privacy and Security Label	128
8.4.2	Tool to Generate the Label	129
8.5	Future Directions to Enhance the Label	129
8.5.1	Design Elements of the Label	129
8.5.2	Actual Behavior vs. Stated Behavior	132
8.5.3	Monetary Valuation of the Label	132
8.5.4	Labels from the Systems Perspective	132
8.6	Path to Label Adoption	134
	Bibliography	137
	Appendix A Survey from "Privacy Expectation and Preferences..."	163
A.1	Sample Survey Scenario	163
A.2	Summary Questions	165

A.3	IUIPC Questions	165
A.4	Demographic Questions	167
Appendix B	Survey from “The Influence of Friends and Experts...”	169
B.1	Survey Scenarios	169
B.2	Sample Survey Questions	170
B.2.1	Questions Posed at the End of Each Scenario	170
B.2.2	Questions Posed at the End of Nine Scenarios	172
B.2.3	Demographic Questions	175
Appendix C	Interview Script and Codebook from “Exploring How Privacy and Security...”	177
C.1	Screening Survey Questions	177
C.2	Interview Questions	178
C.2.1	Questions about Electronic Devices	178
C.2.2	Questions about IoT Devices	178
C.2.3	Questions about Label Evaluation	179
C.3	Supplementary Survey Questions	180
C.4	Codebook	182
Appendix D	Interview Scripts, Surveys, and Codebook from “Ask the Experts: What...”	185
D.1	Interview Questions with Privacy and Security Experts	189
D.2	Questions Asked from Privacy and Security Experts on the First Survey	190
D.3	Questions Asked from Privacy and Security Experts on the Second Survey	190
D.4	Interview Questions with IoT Consumers	192
D.4.1	Questions on Risk Communication in Comparative Purchase Process	192
D.4.2	Questions on Information Comparison in Non-Comparative Purchase Process	193
D.4.3	Questions on Risk Communication in Non-Comparative Purchase Process	193
D.4.4	Questions on Information Comparison and Risk Communication of the Secondary Layer	193
D.4.5	Questions on Format and Layout Considerations	193
D.4.6	Questions on Purchase Behavior Scenarios	194
Appendix E	Survey Questions from “Which Privacy and Security Attributes...”	195
E.1	Survey Questions	195
E.1.1	Device-Related Questions	195
E.1.2	Label-Related Questions	196
E.1.3	Additional Attributes	198
E.1.4	Functionality Perception	199
E.1.5	Demographic Questions	199
E.1.6	Consumer Explanations for Attribute-Value Pairs	200

List of Figures

3.1	Relation between factors and comfort level	19
4.1	Extent of reported influence	50
5.1	Prototype label for security camera	61
5.2	Qualitative terminology	62
5.3	Impact of privacy and security on IoT purchase decision	69
6.1	Experts and consumers study flow	75
6.2	Participant in comparative purchase process	79
6.3	Primary layer of our designed IoT label	91
6.4	Secondary layer of our designed IoT label	92
7.1	Probability of increase in risk perception	107
7.2	Probability of increase in willingness to purchase	110
7.3	Risk perception and willingness to purchase scatter plots	112
8.1	IoT security label proposed by the UK government	126
8.2	Finnish IoT security badge	127
8.3	Primary layer of our designed IoT label for Ring Doorbell	130
8.4	Secondary layer of our designed IoT label for Ring Doorbell	131
8.5	Tool to generate the label	133

List of Tables

- 3.1 Factors varied between vignette scenarios 17
- 3.2 SOUPS’17 demographic information 19
- 3.3 SOUPS’17 qualitative codebook 21
- 3.4 GLMM describing comfort level 25
- 3.5 Classifiers to predict comfort level 27
- 3.6 GLMM describing desire to allow/deny 28
- 3.7 Classifiers to predict allow/deny 29
- 3.8 GLMM describing every-time notification 31
- 3.9 GLMM describing once-in-a-while notification 32
- 3.10 GLMM describing first-time-only notification 33

- 4.1 CSCW’18 data analysis factors 42
- 4.2 CSCW’18 qualitative codebook 44
- 4.3 CSCW’18 demographic information 45
- 4.4 Decision making response time summary statistics 46
- 4.5 Differences in allowing IoT data collections 48
- 4.6 GLMM describing following social cues 49

- 5.1 CHI’19 demographic information 64
- 5.2 Purchase behavior categories 65
- 5.3 Quantifying the impact of privacy and security on IoT purchase behaviors 70

- 6.1 S&P experts’ demographic 82
- 6.2 S&P consumers demographic 83

- 7.1 Attribute-value pairs tested in the survey 100
- 7.2 Chapter 7 demographic information 104
- 7.3 Risk perception summary statistics 105
- 7.4 Willingness to purchase summary statistics 106
- 7.5 GLMM describing risk perception 108
- 7.6 GLMM describing willingness to purchase 109
- 7.7 Metrics used to assess risk perception 116

- C.1 CHI’19 qualitative codebook 184
- D.1 Experts’ arguments 189
- E.1 Consumer explanations for attribute-value pairs 200

Chapter 1

Introduction

The Internet of Things (IoT), composed of network-connected physical objects, is growing rapidly. The devices that make up the IoT vary greatly in form and purpose, from sensors that people voluntarily carry on their wrists, to network-connected thermostats, to street lights that count the number of people who pass by. While these devices provide numerous benefits to consumers and businesses [350], their pervasive ability to collect, store, and transfer information about people’s private lives gives rise to significant privacy, security and safety challenges [302, 359, 365]. For example, the Google Home smart speaker was found to record users’ voices [60] and viewing habits without their knowledge [246]. Furthermore, the popular press has reported on Amazon employees listening to audio files recorded by Echo devices [96] and Google failing to mention that its Nest Hub has an integrated microphone [208]. Less prominent IoT manufacturers have also failed to disclose whether they share users’ data with government agencies [348].

Given the scope of data collection and its potential consequences, people find specifying their privacy preferences and making privacy decisions regarding IoT devices to be overwhelming [324]. A key hurdle to people being able to make informed decisions is due to the difficulty of obtaining information about the privacy and security practices of IoT devices in the first place [117, 237]. These challenges become more burdensome as the number of IoT-related privacy decisions increase [92]. Therefore, to fully realize the potential of IoT, individuals need to be sufficiently knowledgeable and aware of devices’ sensing capabilities and privacy and security practices to make informed decisions and that requires nuanced understanding of societal norms and context, as well as individual needs [262, 288].

The objective of this thesis is to establish a thorough understanding of how users make privacy-related decisions when interacting with IoT devices, combine the obtained knowledge with experts’ insights to develop a privacy and security label for IoT devices, and finally evaluate its usability and risk communication to effectively inform consumers’ IoT-related purchase decision making.

In Chapter 2, we summarize the related work and highlight how our research would contribute to the privacy and IoT literature. We then talk about the first step to help users make informed IoT-related privacy and security decisions, which is to understand their privacy and security preferences and concerns. Chapter 3 is devoted to the description of an experiment for statistically modeling privacy concerns and preferences related to a diverse set of IoT data col-

lection scenarios. In this chapter, we report on a 1,007-participant vignette study focused on privacy expectations and preferences as they pertain to a set of 380 IoT data collection and usage scenarios. In this study, our participants were presented with 14 scenarios that varied across eight categorical factors, including the type of collected data (e.g., location, biometrics, temperature), how the data is used (e.g., whether it is shared, and for what purpose), and other attributes such as the data retention period.

Our findings show that privacy preferences are diverse and context dependent; participants were more comfortable with data being collected in public settings rather than in private places, and are more likely to consent to data being collected for uses they find beneficial. They are less comfortable with the collection of biometrics (e.g., fingerprints) than environmental data (e.g., room temperature, physical presence). We also found that participants are more likely to desire to be notified about data practices that they are uncomfortable with. Finally, our study suggests that after observing an individual's decisions in just three data-collection scenarios, it is possible to predict their preferences for the remaining scenarios, with our model achieving an average accuracy of around 81%. The prediction power

As increasingly large numbers of IoT devices collect personal data, users face more privacy decisions, which could easily overwhelm them [324]. One way to alleviate the burden of decision making is to provide informative social cues about how others have decided in similar IoT data collection scenarios. Social influence has been demonstrated to have a strong impact on people's decision making in many domains [6, 68, 312] as people look at others' behaviors and opinions to inform and improve their own judgments [12, 125]. To better understand which social cues are relevant and whose recommendations people are more likely to follow, in Chapter 4, we report on a Mechanical Turk (MTurk) study with 1000 participants, who were presented with nine IoT data-collection scenarios. Some participants were shown the percentage of privacy and security experts or friends who allowed data collection in each scenario, while other participants were provided with no social cues. At the conclusion of each scenario, participants were asked whether they would allow the described data collection.

Our results help explain the circumstances under which users are more or less likely to be swayed by the reported behavior of others in similar scenarios. For example, our results indicate that when friends denied data collection, participants were more influenced than when friends allowed data collection. On the other hand, participants were more influenced by experts when they allowed data collection. We also observed that influence could get stronger or wear off when participants were exposed to a sequence of scenarios. For example, when experts and friends repeatedly allowed data collection in scenarios with clear risk or denied it in scenarios with clear benefits, participants were less likely to be influenced by them in subsequent scenarios.

As mentioned above, in Chapters 3 and 4, we explore people's privacy preferences, expectations, and concerns related to common IoT data collection scenarios to facilitate informed privacy decision making without being overwhelmed. Asking participants to specify their concerns and preferences related to a diverse set of IoT data collection scenarios helped us obtain a broad view on privacy attitudes and identify the most effective factors that explain people's privacy-related behaviors and decisions. For the rest of this thesis, we focus on a key decision that people are increasingly involved in these days: purchasing IoT devices.

Sales of IoT devices are skyrocketing [144] and various 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 [65, 196]. In Chapter 5, we describe an experiment we conducted to study the significance of privacy and security in consumers' IoT purchase decision making. 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, opinions shared by 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 a prototype of our privacy and security label. Almost all found it to be accessible and useful, reporting that it encouraged them to incorporate privacy and security in their IoT purchase decisions.

Knowing how interested consumers are in having usable and informative privacy and security labels when purchasing IoT devices, we continued improving our prototype label design. While legislators have proposed adding succinct, consumer accessible, labels, they do not provide guidance on the content of these labels. Therefore, the next question we asked ourselves was what information should be included on these labels to convey the most important privacy and security attributes of IoT devices. We answer this question in Chapter 6, where we report on the results of a series of interviews and surveys with privacy and security experts, as well as consumers. In this series of expert and consumer studies, we explore and test the design space of the content to include on an IoT privacy and security label.

We conduct an expert elicitation study by following a three-round Delphi process with 22 privacy and security experts to identify the factors that experts believed are important for consumers when comparing the privacy and security of IoT devices to inform their purchase decisions. Based on how critical experts believed each factor is in conveying risk to consumers, we distributed these factors across two layers—a primary layer to display on the product package itself or prominently on a website, and a secondary layer available online through a web link or a QR code. We report on the experts' rationale and arguments used to support their choice of factors. Moreover, to study how consumers would perceive the privacy and security information specified by experts, we conducted a series of semi-structured interviews with 15 participants who had purchased at least one IoT device (smart home device or wearable). Based on the results of our expert elicitation and consumer studies, we propose a prototype privacy and security label to help consumers make more informed IoT-related purchase decisions.

The effectiveness of any proposed privacy and security label depends on how well the presented information conveys risks to consumers and potentially influences their willingness to purchase the IoT device. In Chapter 7, we discuss the study we conducted with 1,371 MTurk participants to quantify the effectiveness of each of the privacy and security attribute-value pairs we proposed to include on an IoT label along two key dimensions: ability to convey risk to consumers and the impact on their willingness to purchase the IoT device. We found that the aforementioned values intended to communicate increased risk were generally perceived that way by participants. For example, we found that consumers perceived more risk when a label conveyed that their data would be shared with multiple parties than when it would be shared only with the device manufacturer, and that consumers were more willing to purchase devices when they knew that their data would not be retained or shared with others. However, participants' risk perceptions did not always align with their willingness to purchase, sometimes due to usability

concerns. Based on our findings, we propose actionable recommendations on how to further revise our proposed label to more effectively present privacy and security attributes on an IoT label to better communicate risk to consumers. We conclude this thesis by providing final thoughts and future directions in Chapter 8.

All the interview and survey projects conducted as parts of this thesis have been reviewed and approved by Carnegie Mellon University's Institutional Review Board (IRB). All participants provided their informed consent to participate in the surveys and interviews, to have their voices audio recorded, and to have the recordings transcribed by a third-party transcription service. 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.

The interview scripts, survey questions, and qualitative codebooks that were designed and used in the aforementioned studies are all provided in the appendix.

Chapter 2

Background and Related Work

In this chapter, we outline the related work in four sections. To understand people’s privacy preferences related to IoT data collection scenarios, we first highlight the factors that have been shown to be effective in explaining and predicting privacy concerns and attitudes (Section 2.1). Next, we discuss previous research on decision-making processes focusing on consumers’ purchase behavior (Section 2.2). We then provide a background on how consumers perceive risks and what factors impact their risk perception (Section 2.3). Finally, in Section 2.4, we discuss risk communication and product labels to gain insight into how to more effectively inform consumers’ purchase processes.

2.1 Privacy Decision Making

New methods of data collection in the IoT have led to new privacy challenges. Some of these challenges include obtaining consent for data collection; allowing users to control, customize, and choose the data they share; and ensuring the use of collected data is limited to the stated purpose [278]. These challenges are made more difficult by the increased potential for misuse of personal information in the IoT domain. This stems from the pervasive tracking of habits, behaviors, and locations over a long period of time. There are new risks to personal safety introduced by IoT systems [44, 76].

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 [157]. Moreover, experts warn about IoT device security vulnerabilities [21, 154, 299] that could allow an attacker to control a device or collect private data [83, 97, 98, 272, 359]. These vulnerabilities include insecure authentication mechanisms [274], transmitting unencrypted data [33, 353], and failure to promptly patch known bugs [155]. In addition, some devices collect sensitive information and transmit it to the device manufacturer or other parties, raising privacy concerns [20, 180, 227].

Researchers have shown that the extent of these concerns can be explained and predicted based on various factors, such as the type of data collected, the purpose of data collection, and the retention of collected data. We start this section by discussing the factors that have been shown to be effective in explaining people’s privacy concerns and preferences and we then talk about ways to predict such preferences.

2.1.1 Factors Impacting Privacy Preferences and Concerns

Prior studies outside of the IoT context have examined different factors that can impact individuals' willingness to share information based on measures of comfort with data collection. Bilogrevic et al. found that the comfort levels associated with sharing data are highly dependent on the specific type of data and the sharing context (e.g. search engines, social networks, or online shopping sites) [46]. Leon et al. tested whether data retention, access to collected information, and the scope of use affected willingness to share data for online behavioral advertising purposes. Individuals were more willing to share certain types of data if it had a retention period of one day, but for periods longer than one week, individuals were less likely to be willing to share [206].

Other work has focused on privacy preferences related to mobile devices and applications. Lin et al. evaluated individuals' perceptions of requests to access privacy-sensitive resources (e.g., sensors) on mobile devices. They found that both individual expectations of what an app does and the purpose for which an app requests access to sensitive resources impacts their privacy decisions [210]. In order to better understand people's attitudes toward sharing their location in mobile applications, Sadeh et al. built a system that enabled mobile device users to select and limit with whom they want to share their location. They concluded that increasing people's awareness has a critical role in helping them define more precise policies for protecting their privacy [289]. Tsai et al. studied the impact of giving feedback to mobile device users. Their study informed participants about who their data is being shared with, and when the data was shared. The goal was also to help people manage their privacy on a location sharing application. They reported that when people get adequate feedback, they are more willing to share data and more comfortable with sharing their location [326].

In the context of IoT, studies have evaluated several factors that may impact privacy concerns related to IoT data collection. Lederer et al. studied the relative importance of two factors—the entity collecting data, and the situation in which it is being collected—for determining users' privacy preferences in ubiquitous computing settings. Their results indicate that individuals base their privacy decisions on who is collecting their data, rather than the context in which it is being collected [200]. Lee and Kobsa tested five factors related to the context of data collection in two separate studies and found that individuals generally thought that monitoring in personal spaces was unacceptable, along with monitoring by an unknown entity or the government. Their results also indicate that photo and video monitoring may cause some privacy concern regardless of context [202, 203]. Other small, qualitative studies have focused on individuals' privacy preferences related to wearable sensors. These studies revealed that people demand ownership of the data they produce, and that privacy concerns vary depending on factors including retention time and the perceived value of the data collected [38, 187].

According to Bhaskar et al., a major limitation of prior work studying privacy in IoT environments is that studies typically focus on a single environment in which IoT sensing is occurring [44]. Our work, described in Chapter 3, attempts to address this shortcoming by identifying privacy concerns in multiple heterogeneous scenarios which employ different types of data collection. This way, our methodology can determine which factors have the greatest impact on measures of individuals' comfort with data collection. The results can inform the design of privacy-enabling solutions appropriate to the variety of contexts we have studied. Furthermore,

our study aims to expand beyond prior work in this area by identifying privacy concerns individuals have in data collection scenarios which are not obviously aligned with specific privacy risks.

So far, we discussed the factors that impact people’s privacy decision making related to IoT data collection scenarios. However, decision making does not always happen in a vacuum as individuals rely on others’ judgements and adapt their perceptions based on them. Therefore, it is imperative to understand how social cues impact people’s preferences.

2.1.2 Decision Making under Social Influence

Researchers have studied the power of social influence on individual opinions since the early 20th century. Jenness first studied conformity in 1932 and ran an experiment to understand how human estimation changes based on the influence of the majority; he observed that almost all the participants changed their opinions to be close to the group estimate [174]. In Sherif’s well-known Autokinetic Effect experiment, when participants were unsure, they relied on information from others to form their own opinion [298]. In another classic psychology experiment, Asch studied the extent to which majority opinion could affect individual decisions and judgments. In his famous conformity experiment, he asked participants in a group setting to perform a judgment task, in which they had to guess the closest line to the target line. He found that an individual would conform to the majority’s opinion even when the correct answer was obvious. He showed that social influence can make people question their decision when it is different from the majority or simply exhibit public conformity to avoid contradicting group norms [22]. Kelman identified three different types of conformity: *compliance*, *internalization*, and *identification* [185]. *Compliance*, is conformance to meet a specific requirement or to avoid a specific punishment. In contrast, *internalization* occurs when individuals conform to something they believe in and consider a useful solution to their problem. Participants in Sherif’s conformity experiment mostly fell into this category. *Identification*, applies to individuals who conform to fulfill their desire for a relationship with another person or group.

Deutsch and Gerard distinguish *normative social influence* and *informational social influence* [105]. *Normative social influence* occurs mostly when individuals want to fit in with a group, a famous example of which is Asch’s aforementioned line experiment. *Informational social influence* occurs when people seek information and guidance. For example, in Sherif’s Autokinetic experiment, when participants were unsure about the correct answer, they observed other members of the group to inform their own decisions. Informational influence serves as a “cognitive repair” that lessens the harm of depending too much on one’s own judgments [161].

In Chapter 4, we report on a large-scale experiment studying the impact of *indirect informational social influence* in the context of privacy-related decision making and identified factors that can predict individual responses to social influence. In this study, we presented our participants with vignettes that described IoT data collection scenarios. The vignette methodology, which is a technique to elicit opinions and attitudes of individuals by analyzing their responses to different scenarios [37], is particularly useful for determining significant factors and explaining the extent of their impact [7].

The factors that we studied were identified in the psychology literature as relevant for understanding the impact of social influence on decision making, but had not previously been examined

in the context of IoT privacy. We tested the significance of three factors by varying them in the vignettes: expertise [160, 175], consensus level [232], and opinion difference [244].

The literature shows that people heed advice from experts more than advice from less informed individuals. In our study, we tested the impact of expertise by presenting participants with social cues from either privacy experts or friends. In addition, research shows that the level of consensus influences an individual's decision making. For instance, Martin et al. found that people were more influenced when presented with a stronger consensus than when presented with a weaker consensus [232]. In another experiment, Mackie tested the impact of 64% versus 82% of consensus level on decision making [221]. Inspired by these experiments, we tested the impact of two levels of consensus (85% vs. 65%) on participants making privacy-related decisions regarding IoT data collections. We studied how close the participants' initial opinions were to influencers' opinions when making privacy-related decisions. Meshi et al. used the term *opinion difference* to describe this factor. With their fMRI experiment, they found that advice utilization will increase as the opinion difference becomes smaller [244].

Morton and Sasse [252] classified the information-seeking behavior of participants who try to decide whether or not they want to use a service. They found that only 15% of their 58 participants are “crowd followers” who are heavily influenced by “environmental cues” such as the advice of others or media reports when making privacy-relevant decisions. They identified four other groups—information controllers, security concerned, benefit seekers, and those looking for organizational assurance—that all have different demands and interests when making privacy decisions.

Mendel and Toch [243] studied the phenomenon of social influence on privacy behavior in online social networks. We confirm some of their results in the IoT privacy setting. For instance, user's susceptibility to privacy influence depends on his or her privacy knowledge and self-efficacy.

Several other projects have investigated the impact of social influence in privacy or security settings. Patil et al. [276] studied how the preferences of an individual's social circle affects the privacy settings that are selected when using an instant messaging application. They found that this additional information provided useful guidance, but this influence was secondary to the privacy-sensitivity of the setting itself. Besmer et al. [43] studied access control settings for third-party social network applications, measuring the influence of information about the percentage of other users who shared information with such applications. They found that this information could impact user decisions, but only if the cues were sufficiently strong. Balestra et al. [31] studied the impact of exposure to social annotations on privacy consent for a genomics application. These annotations had the general effect of making users feel more informed, but also less confident in their understanding of the application and less trusting in the institution soliciting the consent. Das et al. [93] studied how social influence affects Facebook security settings. They found that many friends adopting a particular feature would influence users towards adoption. Conversely, few friends adopting a feature may bias users away from adoption, which is viewed as not ideal for security features where the goal is to encourage adoption as much as possible.

Social influence has also been examined outside of privacy and security contexts. For example, collaborative filtering systems for product recommendations [192] and reputation management systems, such as the ones used by eBay [283], are examples of social influence in decision making.

2.1.3 Predicting Privacy Preferences

Past research focused on segmenting users based on their privacy preferences has led to heavily-cited differences between privacy fundamentalists or highly concerned, pragmatists or moderates, and marginally concerned or minimalists [194]. However, the narrative that consumers make rational decisions on privacy matters has been challenged [108]. Current approaches to classify users based on their privacy concerns have, therefore, concentrated on finding other indicators for privacy behavior such as knowledge or motivation [109], or simply on using previous choices to predict future decisions and generate recommendations [190, 213].

Prior work has shown that privacy preferences can be inferred by segmenting collections of individuals based on profiles. These profiles represent clusters of different individuals and their privacy decisions. In the mobile app privacy domain, Lin et al. and Liu et al. demonstrated that a small number of profiles may be capable of predicting individuals' decisions to allow, deny, or be prompted for app permissions with a high level of accuracy [211, 215]. In IoT data collection scenarios, Lee and Kobsa were able to identify four clusters of participants with distinctive privacy preferences. These clusters were used to predict their study participants' decision to allow or deny monitoring in a particular IoT context with 77% accuracy [203]. In our work, discussed in Chapter 3, we incorporate additional factors into a larger scale study, using similar techniques to make predictions with the goal of achieving improved prediction accuracy relative to prior work.

2.2 Purchase Decision Making

Despite privacy and security concerns, consumers are still purchasing IoT devices, mostly due to their perceived convenience and features [69]. This is sometimes referred to as a “privacy paradox,” due to the discrepancy between privacy concerns and actions taken to mitigate those concerns [34, 265]. Previous research has identified various explanations for the discrepancy between reported privacy attitudes and actual behaviors. A few of which are biases and heuristics [1, 53], behavioral manipulation and skewing [234, 344], framing effect [4, 351], and inertia and friction [25, 240].

Another mentioned explanation for the privacy paradox is misunderstanding and lack of information [328, 329], which is likely to happen as consumers are provided with little, or often no, privacy and security information about IoT devices prior to purchase [48, 49, 102]. This prevents consumers from making informed IoT-related purchase decisions and increases the risk of privacy and security vulnerabilities, which may lead to high-profile and large-scale attacks targeting IoT devices [19].

We start this section by providing background on purchase processes. We identify the factors that have been shown to impact consumers' purchase decisions in previous research.

2.2.1 Purchase Process and Willingness to Purchase

Purchase behavior is defined as the set of decisions people make and the actions they take when buying and using a product [23]. Purchasing comprises seven stages: need recognition, infor-

mation search, pre-purchase alternatives evaluation, purchase, consumption, post-consumption evaluation, and divestment [47]. Pre-purchase behavior involves deciding on what to buy and when to buy a product [292], 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 [17].

When it is not possible to observe consumers' actual purchase behavior, we ask them about their willingness to purchase. Previous work has shown that willingness to purchase is an indicator of actual purchase behavior [280] and has been shown to have a high correlation with it [9, 269, 360].

2.2.2 Factors Impacting Purchase Decision

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 [179, 212, 220, 290]. The perceived quality of a product has been identified as the main driver of consumers' purchase intentions [271]. Digital and social media have also been shown to impact consumers' purchase behaviors [281]. In addition, word of mouth and reviews have been identified as influential factors [178, 284, 364]. Trust is another factor, which has been shown to impact consumers' purchase behaviors, especially under uncertainty when doing comparison shopping [91, 238]. This trust can be developed through a number of factors such as the size and the reputation of the company [173]. Company reputation has been shown to be closely related to the familiarity with the brand [145].

Studies have found that people are concerned about the privacy of their personal data when making online purchases [81, 169, 328]. Availability of privacy information has been shown to impact consumers' purchase decisions. For example, 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 [325]. In another study, Kelly et al. found that adding concise privacy information to a mobile app store can impact users' app-selection decisions [183]. We conducted an interview study to understand the importance of privacy and security information in consumers' IoT-related purchase process. This study is presented in Chapter 5.

2.3 Risk Perception

Perceived risk is a subjective assessment of the likelihood of a specific event happening and concern about its consequences [306]. In the 1960s, risk perception research started by focusing on risk comparison [310]. Later, Starr found that society is more willing to accept risks that are perceived to provide benefits [313]. Perceived risk has been found comparable to the real risk [342] and people tend to base their decisions on the perceived risk rather than the actual risk [294, 327]. In 1978, Fischhoff et al. [130] identified nine dimensions to measure the extent of perceived risk and found that the dimensions of dread and novelty best explain risk perception [130].

In the context of privacy and security, Gerber et al. found that lay users perceive abstract and specific privacy-related scenarios differently [147]. Abstract scenarios were evaluated as

more likely but less severe, while specific scenarios were deemed rare, but more concerning. Researchers have shown that users' lack of risk awareness and knowledge about how their data might be used [10, 41, 158, 186, 361] influences their risk judgement [3, 28, 142]. Skirpan et al. found identity theft, account breach, and job loss as the top three rated risk scenarios related to emerging technologies [307].

Researchers have examined how people perceive risks of smart devices. Wieneke et al. conducted a qualitative study on how privacy affects decision making related to wearable devices. They found that users' lack of awareness impacts their risk perception and they also observed a disparity between risk perception and behaviors [349]. Their findings are aligned with other work in this space [361, 362].

Consumers' risk perception often differs from that of experts [356]. Therefore, in Chapter 7, we report on a large-scale user study, in which we measured the significance of IoT privacy and security attributes identified in our expert study (Chapter 6), along with factors previously found to explain risk perception, including risk target [259, 295, 304], familiarity with the technology [122, 135, 143], attitudes and concerns [132, 304, 305], and order effects [57]. In our study, we considered the recipient of the device to evaluate the risk target, checked whether participants had the device to gauge familiarity with the technology, and varied the type of device to gauge the impact of concerns related to the type of collected data.

2.4 Risk Communication

Researchers have identified three types of activities that can lead to the development of effective risk communication: mental model analysis, calibration analysis, and value-of-information analysis. In mental model analysis, lay users' mental models are studied in order to identify how they perceive risk and what information could help them make more informed risk-related decisions. Calibration analysis identifies users' most common misunderstandings about risk [51]. Value-of-information analysis systematically identifies the information that will most effectively impact users' risk perception and decision making [129].

One method of risk communication is by using labels, which has been shown to be effective in various domains. In the rest of this section, we will focus on labels as one of the methods to convey risk to consumers.

2.4.1 Labels

Labels are a common approach in contexts such as food [123] and energy ratings [137, 335] to effectively communicate important information to consumers. The literature on nutrition labels has identified several factors that can change the impact of label on consumers' purchase behavior, indicating that the label may not be as effective for every consumer. Some of these factors are label formatting, label wording as well as consumers' age, gender, and income [18, 61, 72, 250, 346]. Despite the influence of these factors, food nutrition labels have been shown to significantly inform consumers' purchase decisions [247, 258].

In the privacy context, researchers have found that posting concise privacy "nutrition labels" on websites increases the speed and accuracy of users' information seeking compared to finding

information in privacy policies [182]. Likewise “privacy facts” checklists in app stores [183] impact users’ app download decisions.

In the context of IoT, 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 [254]. However, it is not designed as a label and is not attached or linked to devices in a store. In addition, when manufacturers do not disclose some information, the guide for a product may be incomplete. Moreover, as far as we know, the buyers guide has not undergone user testing. In Chapter 5, we discuss an interview study we conducted with IoT consumers and found that participants would like to consider privacy and security in their purchase process, if such information was made available to them in a label format.

International Labeling Schemes

Acknowledging the significance of labels in increasing consumers’ awareness when purchasing smart devices, governments have started developing labeling schemes for IoT devices. Governments of the UK [100], Finland [128], and Singapore [321] are the forerunners in labeling their smart devices. All the current labeling proposals are expected to be voluntary to help IoT manufacturers and the market adjust to the guidelines. Unlike our proposed IoT label, the main focus of these proposed labeling schemes is the security practices of smart devices, with a little attention paid to their privacy attributes. In Section 8, we will discuss the attributes currently mentioned in the international labeling proposals.

2.4.2 Privacy and Security Guidelines and Best Practices

Policymakers [102, 103, 104, 121, 124, 209, 228, 267], industry groups [316, 318], and certification bodies [56] have expressed interest in having privacy and security labels for IoT devices. However, the format of such labels and the information they should contain has not been widely discussed.

Tanczer et al. [318] conducted an extensive literature review on publicly available reports from industry associations and international organizations on their security-related proposals and recommendations for consumer IoT devices. They reviewed 17 industry reports (including from Intel, HP, and Consumer Technology Association) and policy reports (including from European Commission, International Organization for Standardization (ISO), and Alliance for the Internet of Things Innovation (AIOTI)). This review observed 19 overarching principles related to security best practices that were referred to at least twice in these reports. The most common principles (mentioned in at least 10 reports) were strong authentication by default, reliable and cryptographically signed security updates, encryption by default, and compliance and risk assessment. Some of the other principles were related to physical security, vulnerability reporting and disclosures, and secure device boot. The security factors that we synthesized based on our expert elicitation study covered all of the most frequent security principles mentioned in this literature review [318].

Tanczer et al. concluded that in general, the industry acknowledges the importance of selling safe and secure IoT devices in the market and would like to work alongside the government

as part of their efforts. However, they are more interested in self-regulation than in government interventions [318]. For example, companies can self-certify their IoT devices using a framework developed by IoT Security Foundation (IoTSF) that specifies five levels of compliance [171].

A recent UK government report argued against self-regulation, noting the lack of incentives for IoT companies to adhere to security best practices when designing IoT products. The report recommended that the government mandate specific requirements for IoT devices as a mechanism to improve the security of consumer IoT products [103]. These requirements are no default password, availability of a vulnerability disclosure program, and security updates. These recommended requirements all are included in our proposed privacy and security label.

Notably, all the reviewed reports above focus on devices' security mechanisms with few references to data privacy considerations. As consumers are concerned about both the privacy and security of their devices, our proposed privacy and security label includes both privacy policies and security mechanisms of an IoT device. In Chapter 6, we discuss our project designing an informative IoT privacy and security label by interviewing and surveying experts. We specify 47 critical privacy and security factors and propose a layered label to present those factors, primarily based on experts' opinions.

Chapter 3

Privacy Expectations and Preferences toward IoT Data Collections

With the rapid deployment of Internet of Things (IoT) technologies and the variety of ways in which IoT-connected sensors collect and use personal data, there is a need for transparency, control, and tools to ensure that individuals' privacy requirements are met. Studying people's privacy requirements needs a nuanced understanding of societal norms and context, as well as individuals' needs [263, 287]. For example, most people tacitly accept being recorded on cameras and CCTV outdoors in public spaces, but express disdain for installing video surveillance systems inside the walls of their homes. As more complex IoT scenarios become possible, many other factors may play a role in determining individuals' privacy preferences. While some may feel comfortable with their location being tracked for the purpose of traffic prediction, they may consent to tracking only their work commute. Others may consent only if they are assured that their location data is retained and used in an anonymized form.

We conducted a large-scale online vignette study to identify the contribution of different factors related to IoT data collection and use scenarios (such as the type of data, retention time, purpose of data collection, and location of data collection) in promoting or inhibiting individuals' self-professed comfort levels. We also studied the factors that trigger a desire for notifications about data collection. Our research identified which aspects of data collection or use by various IoT devices are most likely to cause discomfort, how realistic participants think these scenarios are, and about which aspects they would like to be made aware.

The results of our study informs the design of more transparent IoT-connected systems—we envision our results can be used to improve privacy notices for IoT devices, and develop more advanced personal privacy assistants [214].

This chapter makes two main contributions. First, we show that individuals' comfort levels in a variety of IoT data collection scenarios are related to specific aspects of that data collection and use. Many of our findings are consistent with observations made in prior work, but our quantitative methodology and the scale of our experiment allows us to understand the ef-

This chapter is a lightly edited version of a paper previously published as: Pardis Emami-Naeini, Sruti Bhagvatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, Norman Sadeh. Privacy Expectations and Preferences in an IoT World. In Proceeding of the 13th Symposium on Usable Privacy and Security (SOUPS), 2017 [115].

fect of individual factors and their relative importance more precisely. Second, leveraging our qualitative and quantitative results, we advance explanations for many of the differences among these factors. We show that whether or not participants think the use of their data is beneficial to them has a profound influence on their comfort level. We also find that participants’ desire for notification is closely related to whether or not they feel comfortable with data collection in a particular scenario.

This chapter is organized as follows. First, we describe the design of our vignette study, and discuss our quantitative and qualitative analysis of our survey data. Next, we present the results of our prediction model, and draw conclusions from the analysis. Finally, we will provide some concluding thoughts based on what we learned from this project.

3.1 Methodology

We conducted a within-subjects survey with 1,014 Amazon Mechanical Turk (MTurk) ¹ workers in order to understand individuals’ privacy preferences. We exposed each participant to 14 different vignettes presenting an IoT data collection and use scenario. Vignettes are “short stories about hypothetical characters in specified circumstances, to whose situation the interviewee is invited to respond,” [126] and have been used in prior work studying varying privacy contexts [230, 231].

Between vignettes, we varied eight factors that we hypothesized could influence individuals’ privacy preferences:

- the type of data collected (`data_type`)
- the location where the data is collected (`location`)
- who benefits from the data collection (`user_benefit`)
- the device that collects the data (`device_type`)
- the purpose of data collection (`purpose`)
- the retention time (`retention`)
- whether the data is shared (`shared`)
- whether additional information could be inferred from the collected data (`inferred`)

Several of these factors have already been shown in prior work to be important to individuals, when presented individually or in combination [38, 187, 200, 202, 203, 206]. Our design allowed these factors to be studied simultaneously, capturing more contextual nuances. In our vignettes, some factors could take on one of many possible levels. For reference, table 3.1 describes the factors and their corresponding levels.

After accepting the MTurk HIT, each study participant was directed to a survey where they were shown 14 different vignettes.

Each vignette introduced the factors being tested in the same order. In each scenario, vignettes began with the location of the data collection and ended with the retention period. The following is an example of a scenario presented to participants:

¹Amazon’s Mechanical Turk <https://www.mturk.com>

Factor	Levels	Description
location	department store; library; workplace; friend's house; home; public restroom	location where the data is collected
data_type	presence; video; specific position; biometric data (e.g., fingerprint, iris, face recognition)	type of data collected
device_type	smart watch; smart phone; camera; presence sensor; temperature sensor; fingerprint scanner; facial recognition system; iris scanner	device that is collecting the data; some devices like smart phones can collect multiple data types
user_benefit	user (e.g., get help in emergency situations); data collector (e.g., downsize staff)	who benefits from the data collection and use
purpose	a specific purpose is mentioned; it is mentioned that participants are not told what the purpose is	purpose of data collection depends on the location, the data, and who is benefiting
retention	forever; until the purpose is satisfied; unspecified; week; year	the duration for which data will be kept
shared	shared (e.g., with law enforcement); no sharing is mentioned	whether the data is shared or not
inferred	inferred (e.g., movement patterns); inferred data is not mentioned	Additional information can be inferred and users can be deanonymized

Table 3.1: Factors varied between vignette scenarios, levels of the factors presented in scenarios, and description of each factor.

You are at **work** and your **smart watch** is keeping track of your **specific position in the building**. Your position is shared with the **device manufacturer** to **determine possible escape routes in the case of an emergency or a hazard**. This data will be kept by the manufacturer **until you leave for the day**.

All factorial combinations of the different levels of each factor produced 126,720 possible scenarios, many of which contained combinations of factors which did not make sense (e.g. a presence sensor taking iris scans for emergency purposes). These scenarios were removed from the set of scenarios shown to participants. From the remaining set, we selected 380 scenarios that could feasibly occur, and ensured that this subset contained scenarios in which each level of each factor was represented. 14 vignettes drawn from these 380 scenarios so as to not overburden them. Randomly selecting subsets of 14 scenarios could have caused interaction effects due to a lack of diversity in each factor (e.g., presenting only one retention time on otherwise diverse scenarios) [26]. To minimize such interaction effects, we carefully selected subsets of vignettes so that every level of every factor was present at least once per subset, with the exception of the factors `device_type`, `purpose`, and `inferred`, which were dependent on other factors such as `location`, `device_type`, and `user_benefit`. In doing so, we divided the list of scenarios into 39 subsets with 14 scenarios each, and presented each participant with vignettes corresponding to one of these 39 subsets. The subsets were not mutually exclusive.

For each scenario, participants were asked how comfortable they were with data collection in that scenario and whether they found the use of data in the scenario to be beneficial

(`user_perceived_benefit`). This factor is different from `user_benefit`, which refers to whether the data collection benefits the participant or the collector and is part of the scenario design; `user_perceived_benefit` refers to the participant’s perception of whether the scenario would be beneficial to them. This question was only asked about scenarios in which a purpose was given; we coded this factor as ‘N/A’ for scenarios without a purpose. We also asked participants whether they would allow the data collection described in the scenario, and how often they would like to be informed about the data collection. Further questions asked how realistic a scenario was (“I think scenarios like this happen today,” “... will happen within 2 years,” and “... will happen within 10 years”) and coded the answers to these three questions as `happening_today`, `within_two_years`, and `within_ten_years`, respectively. These three questions were answered on a five-point Likert scale from “Strongly Disagree” to “Strongly Agree” and were binned into binary categories based on agreement—0 (strongly disagree, disagree) and 1 (strongly agree, agree, neither agree nor disagree). Finally, we asked participants general demographic questions, followed by ten questions from the Internet Users’ Information Privacy Concerns (IUIPC) scale to gauge their level of privacy concern. The IUIPC scale questions focus on concerns about control, awareness, and collection [225]. The complete set of questions asked in our survey is included in Appendix A.

3.1.1 Factors Impacting Preferences

We were interested in learning which factors of data collection and use contributed most significantly to individuals’ comfort and preferences. Thus, we asked questions about how comfortable they were with the given scenario. We also asked if they would allow a specific data collection or not, and how often they would want to be notified about it. Participants’ responses to these questions enabled us to build statistical models that predict the concerns and preferences of the general population, based on our sample. We constructed five statistical models, capturing five dependent variables (DV): comfort level, allow or deny decisions for the data collection, desire to be notified of data collection every time, desire to be notified once in a while, and desire to be notified only the first time. In addition to the eight factors in Table 3.1, we included the factors `user_perceived_benefit`, `happening_today`, `within_two_years`, `within_ten_years`, `gender`, `age`, `income`, and `education`, as well as the three IUIPC scale factors `IUIPC-control`, `IUIPC-awareness`, and `IUIPC-collection`.

We represented income as a quantitative variable based on categories of income ranges, excluding two outliers—participants who reported earning more than \$200,000. We mapped all Likert scale responses to binary categories of 0 and 1, where 1 implies a positive preference, and 0 implies a negative preference. All of the quantitative variables (`income`, `age`, `IUIPC-control`, `IUIPC-awareness`, `IUIPC-collection`) were normalized before analysis to be on the same scale with a mean of 0 and standard deviation of 1.

We did not include two of the eight privacy factors, `device_type` and `purpose`. The device that is collecting the data was mentioned in the vignettes to make them more realistic, but was not considered in the statistical analysis because the device was uniquely determined by the type of data that was collected. The type of data that was collected was considered in the statistical analysis, resulting in a dependency between the two factors. Dependencies of this type between factor levels can lead to inaccurate statistical inferences. To improve the accuracy of

Gender	Age	Education	Income	IUIPC Score
Male 49.2% (49.2%)	Range 18-78	No high school 0.8% (10.9%)	< \$15k 16.4% (11.6%)	<i>Control Factor</i>
Female 50.1% (50.8%)	Mean [SD] 36.1 [10.9]	High school 30.8% (28.8%)	\$15k-\$34k 33.8% (20.5%)	Range 1.33-7
No answer 0.7% (0.0%)	US average 37.9	Associates 9.7% (10%)	\$35k-\$74k 36.1% (29.4%)	Mean [SD] 5.95 [0.90]
		Bachelors 49.0% (48.7%)	\$75k-\$149k 9.3% (26.2%)	<i>Awareness Factor</i>
		Professional 8.5% (1.5%)	\$150k-\$199k 0.9% (6.2%)	Range 1-7
		No answer 1.0% (0.0%)	> \$200k 0.2% (6.1%)	Mean [SD] 6.44 [0.82]
			No answer 3.2% (0.0%)	<i>Collection Factor</i>
				Range 1-7
				Mean [SD] 5.79 [1.11]

Table 3.2: Demographic breakdown of our participants. In the Gender, Education, and Income columns, the numbers in parentheses show the US average, according to census data from 2015.

our results, we excluded them from our statistical analysis. For the same reason, we removed `purpose` as it was not linearly independent from multiple other factors, such as `location` and `user_benefit`. Treating it as an independent factor would have resulted in scenarios that did not make sense contextually. For instance, using `purpose` as an independent factor would have included scenarios which involved collecting fingerprints to downsize staff. To eliminate these nonsensical scenarios from our study, we chose to remove `purpose` from the analysis, instead of the other factors on which it depended.

After removing these two factors, we found one of the subsets of scenarios contained two scenarios that differed only in these two factors. Therefore, for participants who received this subset, we removed the first of the two scenarios’ answers and analyzed the remaining 13 scenarios.

Our models were constructed using generalized linear mixed model (GLMM) regression with a random intercept per participant. GLMM is particularly useful for modeling repeated measures experiments, such as ours, in which participants are presented with multiple parallel scenarios [40].

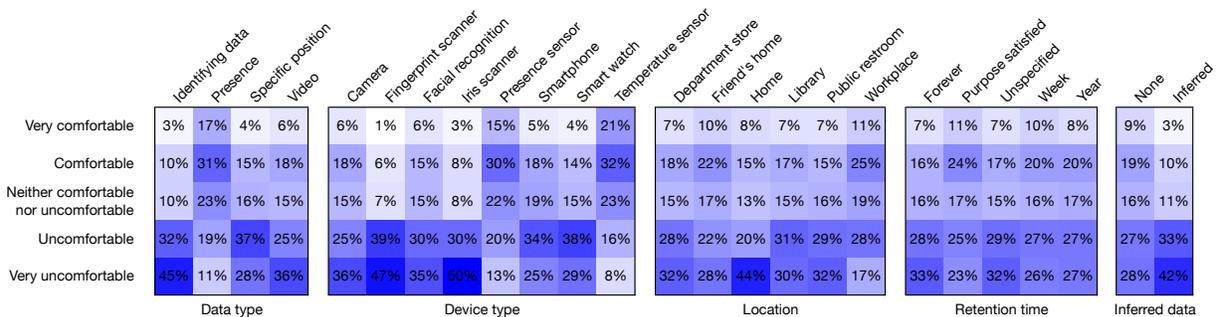


Figure 3.1: Summary statistics showing the relation between various factors and participants’ comfort level. For example 45% of participants were very uncomfortable when the type of data being collected was biometric. Cells with larger numbers are darker in background color.

We performed model selection to find the best combination of factors by using a search algorithm with a backwards elimination approach. For each of our dependent variables, we found the model that best fit the data according to the Bayesian Information Criterion (BIC). We eliminated the variables with the largest p -value in each step of the model selection and continued

the elimination until the BIC reached the global minimum [176]. The model with the lowest BIC best explains the dependent variable.

We present the regression tables for our best models in the Results section 3.2. We used a significance threshold of 0.05 to determine whether or not a factor was significant. Effects and the effect size of a factor level can be interpreted as proportional to the magnitude of the estimate co-efficient. We also defined a baseline for each factor. The regression tables and co-efficients of levels in the model were computed against the corresponding factors' baseline. Some of the baselines were selected based on specific concerns highlighted by our qualitative data, such as `data_type` (baseline = specific position) and `location` (baseline = friend's house). The baselines for other factors were selected based on their alphabetical ordering.

3.1.2 Predicting Preferences

Using the results from the model selection for each dependent variable, we further examined their predictive ability for individuals' preferences. Specifically, in our analysis we focus on predicting:

- an individual's comfort with a specific data collection scenario; and
- an individual's decision to allow or deny a specific data collection instance.

We believe that the ability to predict individuals' preferences or decisions is useful since we can imagine deployment scenarios where a system needs to predict an individual's comfort or decision to allow or deny data collection. In these cases, the system would have more data accumulated over time specific to an individual using the system, and so would likely perform better than the classifiers in our experiments.

Features

For each of the two prediction tasks mentioned above, we used the main factors and interactions from the results of our model selection to predict the two outcomes; comfort level, and the decision to allow or deny.

Continuous features were encoded as-is in the feature vector, while categorical features were encoded as one-hot vectors for each category in the domain of that feature. This means, that each categorical variable was encoded as a vector of binary features where each feature corresponded to the binary value of one of the categories in the original categorical variable. In a one-hot vector, only one value in the whole vector will be 1 at any given time. This is a common way of encoding multi-class categorical features for machine learning tasks. For each categorical variable, the overall feature vector was increased in size by the size of each one-hot vector. For interactions between whole factors, we computed the product of each combination of the values in the one-hot vector and appended this vector of interaction products to the feature vector.

Classifiers

We experimented with various binary classifiers for the allow/deny prediction, and both binary and continuous classifiers for the comfort prediction. For binary classifiers where the outcome

Categories	Tags (Usage)	Examples
Factors (<i>n</i> = 842)	purpose (63%), data (26%), retention (25%), sharing (18%), benefit (17%), location (7%), device (2%)	P880: "It would make me more comfortable knowing where this data was going and how it was going to be used, as well as it being consented."
Whitelist (<i>n</i> = 350)	safety (42%), anonymous_data (40%), personal_benefit (7%), public (7%), common_good (6%), improve_services (6%)	P908: "If they helped to make me safer in some way.", P779: "I'd be fine with data that doesn't identify me.", P121: "That my safety was the reason for it, or saving me money"
Blacklist (<i>n</i> = 474)	biometrics (26%), personal_information (20%), everything (16%), location (13%), private_location (12%), bathroom (9%), video (9%), commercial (8%), government (6%), law_enforcement (5%)	P136: "[...] that they might share the data with other parties [...]. Also, knowing that a retinal or fingerprint scan might be stolen and used to gain access to something else." P415: "The government spying on me in my home, or private corporations using that data to identify me [...], no way."
Information (<i>n</i> = 417)	purpose (66%), retention (35%), sharing (21%), collector (15%), access (13%), data_handling (13%), data_security (5%)	P271: "Knowing exactly what the data is used for, where it is stored, who it is shared with, and when it is collected."
Control (<i>n</i> = 113)	deletion (33%), consent (30%), opt-out (27%), ownership (14%), access (13%), copying (10%)	P913: "Nine times out of ten I won't care and would be happy to allow it, I just want to be informed and have the ability to deny consent should I choose."
Risks (<i>n</i> = 298)	misuse (29%), surveillance (18%), data_security (18%), privacy (16%), tracking (12%), intransparency (8%)	P286: "I don't want my personal information getting into the wrong hands." P47: "I don't like the idea of government organizations being alerted of my location at all times."

Table 3.3: Categories and codes used to code free text answers. Percentages in brackets are the number of times a code was used when the category was coded, multiple codes could be applied per category. Rows on Factor/Whitelist/Information/Control refer to answer to the question “..what would make you uncomfortable with sharing data in such situations?” Blacklist/Risks stem from the answers to the question about discomfort.

is binary, we used logistic regression, support vector machines (SVM), k-Nearest Neighbor, AdaBoost (with various weak base classifiers), and simple neural networks in the form of three-layer multi-layer perceptron (MLP) [277]. For predicting comfort, we also experimented with a continuous version of the comfort level on a scale from 1 to 5, normalized to be between 0 and 1, for which we used linear regression for prediction.

We found the AdaBoost classifier with a logistic regression base classifier (together with L_2 -regularization) to be the best performing, and these are the results we report on. We implemented our classifier and ran experiments using the Scikit-learn Python library [277].

Evaluation Methodology

We tested using two different sizes of the training data for predicting a specific participant’s preferences: 75% of 100% of the answers provided by the remaining participants. In all cases, training data also included the participants’ own answers to three of the scenarios they were asked about; we tested on the remaining 11 scenarios (10 scenarios in the case of the participants mentioned in Section 3.1.1).

When predicting comfort level, we report accuracy in two ways, which differ in how they treat predictions when the participant did not have a preference. In the first approach, we counted any

prediction as correct if the participant’s actual survey response fell in the middle of the Likert scale, i.e., their answer was “Neither Agree nor Disagree.” We did this based on the reasoning that if an individual doesn’t have an explicit preference, then any prediction would be consistent with that preference. In the second approach, we report accuracy by testing only on scenarios for which a participant did not answer neutrally. This measures how many of a participant’s non-neutral preferences can be predicted.

Additionally, for both prediction tasks, we report the results of using a simple majority classifier that classifies each element in the test set as the majority class within the training set.

In each experiment, we randomly selected 50 participants whose answers to predict. We report the accuracy, precision, and recall of the classifier averaged over the 50 participants.

Accuracy is the fraction of predictions that were accurate. Both precision and recall are indicators for measuring the effectiveness of a classifier in predicting positive examples. For predicting comfort, a positive example is a scenario for which the user’s answer falls into the “comfortable” category. For predicting allow/deny decisions, a scenario for which a user answers “Allow” is a positive example. *Precision* is the fraction of positive predictions that are actually correct according to the ground-truth data. *Recall* is the fraction of all the positive ground-truth data that the classifier predicts as positive.

For each participant, we used a form of cross-validation defined as follows:

For $X = 75\%$ or $X = 100\%$ of training data:

- Randomly select 50 participants as targets for prediction.
- For each participant, run 6 different iterations of prediction.
- In each of the 6 iterations, randomly select $X\%$ of training data from the remaining participants and randomly select 3 responses from the total set of scenarios the target was asked about. This data is used for training; testing is done on the remaining scenarios of the target.
- Calculate the average accuracy, precision, and recall scores averaged over 6 iterations each and over the 50 random participants.

We report on the results of our experiments in Sections 3.2.2 and 3.2.3.

3.1.3 Qualitative Analysis of Preferences

We also qualitatively analyzed participants’ responses to the free-response questions they were asked at the end of the survey. The answers were coded with regards to five topics: the factors that were mentioned; whether specific scenarios were described as comfortable or uncomfortable; what the participant wants to be informed about; and what means of control (e.g. access, edit, ability to delete) they request. A codebook was developed from 100 answers and applied to another set of 100 answers by two annotators independently. They reached an inter-annotator agreement of 0.89 (Cohen’s Kappa) for whether a topic was addressed and between 0.67 and 0.72 on the actual tags (e.g., which factor was mentioned). After achieving this accuracy, the remaining answers were divided among the two annotators and coded by one annotator each. A summary of categories and codes and their occurrence is shown in Table 3.3.

3.1.4 Limitations

Our study has limitations common to many user studies including those in the area of privacy. Although the demographic attributes of the participant group are, except for the reported income, close to the US average, Mechanical Turk workers do not reflect the general population. Prior research has shown that Mechanical Turk workers are more privacy-sensitive than the general population [177]. It has also been shown that self reports about privacy preferences often differ from actual behavior. This is referred to as the “privacy paradox” [2, 110]. Our study may be susceptible to this bias because the scenarios were abstract and participants were asked to imagine themselves in situations they may not have encountered. In addition, some of the scenarios in our study were designed to be realistic based on common data collection and use practices that are happening today, while others were designed to be more forward-looking. We decided to have some less-realistic scenarios because we hypothesized that there is a relation between participants’ comfort level about each vignette and their perception of how realistic it is. Nevertheless, participants may have been asked about situations which they are not typically put in, influencing their decisions.

Despite these limitations, presenting a large variety of scenarios to participants allowed us to explore situations that do not currently happen but may be similar to situations that will happen in the future. Since the Internet of Things is still an emerging field, it is not possible to describe situations that are realistic to all participants who may never have had an IoT device or never have faced a situation in which an IoT sensor is collecting data.

3.2 Results

In this section, we describe our participants and present results regarding participants’ comfort level with different data collection scenarios, their decisions to allow or deny data collection, and desire to be notified.

3.2.1 Participants

Our survey was completed by 1,014 MTurk workers. We removed the answers of seven participants because they took less than five minutes to complete the survey, while the average completion time was 16 minutes. This resulted in 1,007 participants whose responses we included in our analyses. Participants were required to be from the United States and have a HIT approval rate of above 95%. Table 3.2 describes participants according to their demographics and privacy concern level. Our participants were slightly better educated and had a higher income than the U.S. average.

3.2.2 Comfort with Data Collection

In our survey, after presenting each scenario we asked: “How would you feel about the data collection in the situation described above if you were given no additional information about the scenario?” We measured participants’ comfort on a five point Likert scale from “Very Comfortable”

to “Very Uncomfortable” with the middle point of “Neither Comfortable Nor Uncomfortable.”

Figure 3.1 shows the general distribution of participants’ comfort across different levels of each factor. Participants were strongly uncomfortable if the scenarios they were asked about had `biometric` as `data_type` (45% strongly uncomfortable), `device_type` as iris scanner (50% strongly uncomfortable), `location` as their home (44% strongly uncomfortable), `retention` as forever (33% strongly uncomfortable), or if other data was `inferred` from the data collection (42% strongly uncomfortable).

Factors Impacting Comfort Level

Using the best model, we ordered the factors based on their contribution to comfort level by looking at the change in BIC when each factor was added to the null model (the model that has no factor other than random intercept for participants). Table 3.4 shows the factors ordered by their effect sizes from the most effective factor (the interaction between the `data_type` and `happening_today`) to the factor with the lowest effect size (`retention`). As shown in the table, not all levels of the factors are statistically significant (p -value < 0.05). A positive estimate (effect size) indicates inclination toward comfort and a negative estimate shows inclination toward discomfort.

Scenarios in which video was being collected and participants thought such data collections are `happening_today` had the greatest positive impact on participant comfort with data collection (p -value < 0.05 , coefficient = 1.39). This is in line with our qualitative results, where we found that 38% of all participants mentioned a specific scenario with which they were comfortable (category “whitelist,” Table 3.3), and from the whitelisted scenarios, 42% mentioned safety, security, or emergency situations as specific purposes for data collection that they would generally approve of. Another 40% of those who whitelisted a scenario were less concerned when anonymous or anonymized data was involved. When an example was given, participants mentioned scenarios involving presence or temperature sensors as ones they would be comfortable with.

Scenarios in which biometric information (e.g., fingerprint, iris image) was being collected and participants thought such data collection is `happening_today`, had the greatest negative impact on participant comfort (p -value < 0.05 , coefficient = -0.89). This is also in line with our qualitative analysis of answers to the question “Keeping in mind the 14 scenarios, what would make you uncomfortable with sharing data in such situations?” In 46% of the answers, participants conveyed one or more specific things that they did not want to happen (coded in category “blacklist,” Table 3.3). Within these answers, the collection of biometric `data_type` was mentioned by 26%.

Based on previous findings [46], we hypothesized that participants would be less comfortable if a scenario included the explicit notice that collected data would be shared with others (`shared`). Consistent with that hypothesis, we found that informing participants that data would be shared with third parties (e.g., with the device manufacturer or law enforcement) caused participants to be less comfortable (p -value < 0.05 , coefficient = -0.68). The qualitative results show that a minority of participants expressed mistrust of or discomfort with sharing with government (6%) and law enforcement agencies (5%).

Within the qualitative responses related to discomfort, we also found explanations of why

Factor	Estimate	Std Err	Z-value	p-value	BIC
<i>data type:happening today</i>					14633
<i>baseline=friend's house:not happening today</i>					
video:happening today	1.39	0.20	6.83	0.00	
biometric:happening today	-0.89	0.15	5.80	0.00	
presence:happening today	0.91	0.18	12.57	0.01	
temperature:happening today	0.95	0.22	4.26	0.00	
<i>data (baseline=specific position)</i>					15843
biometric	-1.45	0.13	-11.12	0.03	
presence	1.42	0.16	8.99	0.00	
temperature	2.50	0.20	12.57	0.00	
video	-0.30	0.19	-1.62	0.11	
<i>user perceive benefit:location</i>					15866
<i>baseline=beneficial:friend's house</i>					
not beneficial:department store	0.00	0.32	0.00	0.99	
purpose unspecified:department store	-0.07	0.24	-0.30	0.76	
not beneficial:house	-0.15	0.48	-0.30	0.76	
purpose unspecified:house	0.05	0.28	0.19	0.85	
not beneficial:library	-0.45	0.33	-1.38	0.00	
purpose unspecified:library	-0.17	0.24	-0.70	0.48	
not beneficial:public restroom	-0.40	0.36	-1.10	0.27	
purpose unspecified:public restroom	-0.48	0.26	-1.85	0.01	
not beneficial:work	-0.49	0.36	-1.38	0.17	
purpose unspecified:work	-0.11	0.24	-0.47	0.63	
<i>being shared:user perceived benefit</i>					15969
<i>baseline=not being shared:beneficial</i>					
being shared:not beneficial	-0.71	0.19	-3.70	0.00	
shared:purpose unspecified	0.37	0.13	2.94	0.02	
<i>user perceived benefit (baseline=beneficial)</i>					16055
not beneficial	-1.88	0.34	-5.60	0.00	
purpose unspecified	-1.30	0.25	-5.26	0.04	
<i>retention:user perceived benefit</i>					16058
<i>baseline =unspecific:not beneficial)</i>					
not deleted:not beneficial	-0.12	0.22	-0.06	0.96	
purpose specific:not beneficial	-0.30	0.28	-1.08	0.28	
week:not beneficial	0.49	0.23	2.11	0.00	
year:not beneficial	0.10	0.24	0.39	0.69	
not deleted:purpose unspecified	-0.43	0.16	-2.69	0.00	
week:purpose unspecified	-0.29	0.16	-1.76	0.07	
year:purpose unspecified	-0.22	0.17	-1.31	0.19	
<i>happening within 2 years (baseline=disagree)</i>					16199
agree	0.96	0.11	9.01	0.00	
<i>happen today (baseline=disagree)</i>					16491
agree	10.98	333.4	0.03	0.97	
<i>location (baseline=friend's house)</i>					17987
library	1.00	0.18	5.54	0.00	
work	0.87	0.18	4.82	0.01	
house	-0.88	0.20	-4.34	0.00	
department store	0.76	0.18	4.24	0.00	
public restroom	0.29	0.19	1.48	0.14	
<i>being shared (baseline=not being shared)</i>					18079
being shared	-0.68	0.09	-7.86	0.00	
<i>IUIPC</i>					
collection	-0.59	0.05	-11.47	0.04	18081
<i>retention (baseline=not specified)</i>					18103
week	0.25	0.11	2.25	0.00	
year	0.16	0.11	1.45	0.14	
purpose specific	0.56	0.15	4.85	0.02	
not deleted	0.10	0.10	0.99	0.32	

Table 3.4: Generalized linear mixed model regression output for the comfort level model. A positive estimate (effect size) indicates inclination toward comfort and a negative estimate shows inclination toward discomfort. Factors are ordered by their contribution: the factor with the lowest BIC contributes most to explaining participants' comfort level.

participants did not want to share their data. About 29% of all participants mentioned some perceived risk, ranging from the fear of identity theft or the use of data for other than the stated purpose (misuse) to a general concern about privacy and surveillance in general. Among those that mentioned a perceived risk, 29% feared that their data could be used in a way that would harm them or put them at a disadvantage. About 18% of these answers explicitly mentioned data security issues and leaks as a cause of concern.

P11: [I'm concerned about] any unique identifiers that could be hacked and then used for identity theft, blackmail, humiliation, etc.

With respect to the location of data collection, most levels had small, positive effect on comfort level. As described above, only scenarios taking place at home had a negative impact on the perceived comfort. Our qualitative results further substantiate this, as participants who mention `location` as a factor that made them comfortable often cited the dichotomy between public and private places. Data collection in private places is described as highly intrusive while data collection in publicly accessible spaces like libraries or stores was described as “ok.” Out of the 474 participants that expressed discomfort with specific scenarios, those that took place in one’s home (12%) and in bathrooms (8%) were most frequently mentioned.

The factor `retention` had the smallest effect size on the results and only short retention times (immediate deletion or storing for a week) had a significant, positive effect on the comfort level. This is in line with the qualitative results were, about 25% of those that mentioned a specific factor in their answers referred to how long their data was stored. Those that explicitly mentioned a time span favored a retention time of less than a week.

Predicting Comfort Level

As explained in Section 3.1, we trained a machine learning model to predict a participant’s comfort based on the significant factors and interactions determined through model selection. The results are shown in Table 3.5.

The classifier achieved an average accuracy of around 81% over 50 different participants when either 100% or 75% of the other participants’ answers are used as training data.

There is a sizable difference in precision and recall depending on whether (1) predictions are counted as correct whenever participants expressed neither a positive nor a negative opinion or (2) scenarios in which participants did not express an opinion are removed from the test data. As per the discussion in Section 3.1.2, both ways of measuring performance are indicative of the utility of using a similar classifier in practice.

Table 3.5 also describes the performance of our simple majority classifier that uses all non-test participants’ answers as training data. These results form a baseline for understanding the performance of the AdaBoost classifier. Although a majority classifier is correct about 70% of the time, AdaBoost additionally correctly predicts more than a third of the predictions that the majority classifier gets wrong.

Classifier	Training	Neutral	Accuracy	Precision	Recall
ABC	100% (1,006)	correct	81.06%	73.86%	83.06%
ABC	100% (1,006)	excluded	77.53%	54.50%	63.49%
ABC	75% (755)	correct	81.79%	71.30%	78.34%
ABC	75% (755)	excluded	77.67%	54.48%	60.77%
SMC	100% (1,006)	correct	72.03%	71.33%	40.92%
SMC	100% (1,006)	excluded	67.96%	0%	0%

Table 3.5: Accuracy, precision, and recall of (1) ABC: the AdaBoost classifier (with logistic regression as the base learner) and (2) the SMC: simple majority classifier, for predicting a user’s comfort level with an instance of data collection. “Training” indicates the fraction (and number) of non-test participants used to train the classifier. “Neutral” indicates whether predictions are always counted as correct if a participant didn’t indicate a preference for that scenario (“correct”) or whether such scenarios are removed from the test set (“excluded”).

3.2.3 Allowing or Denying Data Collection

Factors Impacting Allow/Deny Decisions

We found a set of factors that can explain participants’ response to the question: “If you had the choice, would you allow or deny this data collection?” We again ordered factors with respect to their effect size. The interaction of `data_type` and `location` has the most impact while `shared` has the smallest effect. By looking at the coefficient of the levels within each factor we can claim that participants were most likely to deny data collection in scenarios in which their presence was being collected at their workplace. Also, knowing that the data was being shared had the least effect on their preference to deny a data collection. In this model a positive estimate shows likeliness to deny and a negative estimate shows the likeliness to allow a data collection scenario. The regression results are shown in Table 3.6.

Among the common statistically significant factor levels, the ones that made participants more likely to be comfortable with a data collection also made them more likely to allow the data collection. Many factors were in line between the two models of comfort level and allow/deny such as `data_type`, `location`, `user_perceived_benefit`, `shared`, `retention`, `happening_today`, and `within_two_years`. However, the best model that described participants’ comfort level (Section 3.2.2) was not the same as the best model that described the desire of participants to allow or deny a data collection. For example, we found that the interaction between `data_type` and `location` was the most helpful factor in the allow/deny model, but this factor was shown to be non-significant in explaining the comfort level. This suggests that being comfortable with a specific data collection instance does not automatically mean that someone would allow it to occur, given the choice.

In the free text answers to the questions about what would make them feel comfortable or uncomfortable with data collection, about 11% of all participants mentioned some type of ability to control collection or use as a requirement for comfort, though our scenarios did not include such a feature. Nevertheless, participants expressed interest in a variety of ways to control their personal information. Within the group that mentioned it, 33% wanted to be granted the ability to

Factor	Estimate	Std Err	Z-value	p-value	BIC
<i>data:location</i>					15232
<i>baseline=specific position:friend's house</i>					
biometric:department store	1.58	0.24	6.38	0.01	
presence:department store	1.22	0.37	3.30	0.00	
temperature:department store	1.61	0.55	2.94	0.00	
video: department store	-0.99	0.21	-4.83	0.00	
presence: house	0.42	0.41	1.02	0.31	
temperature: house	0.23	0.42	0.54	0.58	
biometric:library	1.16	0.23	5.01	0.01	
presence:library	1.55	0.37	4.10	0.01	
temperature:library	1.52	0.43	3.52	0.00	
video:library	-0.50	0.20	-2.46	0.00	
presence:public restroom	1.87	0.36	5.11	0.00	
temperature:public restroom	1.54	0.38	3.99	0.00	
video:public restroom	1.36	0.36	3.77	0.00	
presence:work	2.11	0.34	6.10	0.03	
temperature:work	1.66	0.39	4.29	0.00	
<i>being shared:user perceived benefit</i>					15297
<i>baseline=not being shared:beneficial</i>					
being shared:not beneficial	0.62	0.19	3.26	0.00	
shared:purpose unspecified	-0.27	0.12	-2.10	0.04	
<i>retention:user perceived benefit</i>					15352
not deleted:not beneficial	-0.14	0.226	-0.65	0.51	
purpose-specific:not beneficial	0.39	0.248	1.37	0.17	
week:not beneficial	-0.12	0.24	-0.52	0.60	
year:not beneficial	-0.17	0.24	-0.68	0.49	
not deleted:purpose unspecified	0.45	0.16	2.81	0.02	
week:purpose unspecified	0.76	0.16	4.52	0.00	
year:purpose unspecified	0.48	0.17	2.85	0.01	
<i>user perceived benefit (baseline=beneficial)</i>					15374
not beneficial	2.85	0.17	16.38	0.00	
purpose unspecified	1.67	0.17	9.92	0.01	
<i>data:happening today</i>					15525
<i>baseline=friend's house:not happening today</i>					
video:happening today	-1.39	0.22	-6.26	0.00	
biometric:happening today	-0.78	0.16	-4.89	0.00	
presence:happening today	-0.95	0.19	-5.02	0.02	
temperature:happening today	-0.90	0.23	-3.87	0.00	
<i>happening within 2 years:benefit of scenario</i>					15986
<i>baseline=disagree:benefit to company</i>					
agree: purpose unspecified	0.12	0.36	0.34	0.73	
agree:benefit to user	-0.38	0.23	-1.64	0.00	
<i>happening within 2 years (baseline=disagreement)</i>					16751
agreement	-0.72	0.20	-3.70	0.03	
<i>data (baseline=specific position)</i>					16872
biometric	0.01	0.24	0.06	0.95	
presence	-2.87	0.35	-8.01	0.00	
temperature	-3.66	0.37	-9.66	0.00	
video	0.43	0.23	1.82	0.07	
<i>happening today (baseline=disagreement)</i>					17112
agreement	-11.01	349.40	-0.03	0.97	
<i>benefit of scenario (baseline=benefit to company)</i>					18188
benefit to user	-0.46	0.20	-2.30	0.01	
purpose unspecified	-1.17	0.27	-4.34	0.00	
<i>location (baseline=friend's house)</i>					18569
library	-1.87	0.29	-6.34	0.02	
work	-1.96	0.27	-7.34	0.01	
house	0.54	0.35	1.52	0.13	
department store	-1.58	0.29	-5.30	0.00	
public restroom	-1.23	0.29	-4.17	0.04	
<i>retention (baseline=not specified)</i>					18669
week	-0.55	0.11	-4.72	0.02	
year	-0.32	0.11	-2.79	0.00	
purpose-specific	-0.70	0.12	-5.76	0.00	
not deleted	-0.03	0.11	-0.26	0.79	
<i>being shared (baseline=not being shared)</i>					18707
being shared	0.52	0.10	5.41	0.00	

Table 3.6: GLMM Regression Output for the allow-deny model. A positive estimate shows likeliness to deny and a negative estimate shows the likeliness to allow. Factors are ordered by their contribution: the factor with the lowest BIC contributes most to explain participants' desires to allow or deny a data collection.

Classifier	Training	Accuracy	Precision	Recall
ABC	100% (1,006 users)	79.09%	76.79%	82.32%
ABC	75% (755 users)	79.09%	76.79%	82.32%
SMC	100% (1,006 users)	52.58%	0%	0%

Table 3.7: Accuracy, precision, and recall of (1) ABC: the AdaBoost classifier (with logistic regression as the base learner) and (2) SMC: the simple majority classifier, for predicting a user’s decision to allow or deny data collection. “Training” indicates the fraction (and number) of non-test participants used to train the classifier.

delete their data; this would make them feel more comfortable. Another 30% wanted to be asked for consent first, and 27% desired the ability to opt out of the data collection at any time. Multiple participants acknowledged that they would probably not make use of the control options, were they provided.

Predicting Allow/Deny Decisions

Using the significant factors and interactions we determined from the model selection, we trained a machine learning model to predict an individual’s decision to allow or deny data collection. The results are shown in Table 3.7. In this experiment, a prediction is made based on the class (allow or deny) that had the higher probability in the prediction. Averaged over 50 test participants, accuracy ranged from 76% to 80% depending on whether we used most (75%) or all of the other participants’ data during training.

Table 3.7 also describes the results of our simple majority classifier when using all other participant’s answers as part of the training data. Similar to when predicting comfort, we use the results of this experiment as an intuitive baseline for understanding how well a classifier does if it simply uses the most prevalent preference in the training data.

The average accuracy of the majority classifier of barely over 50% shows that participants’ collective preferences were sufficiently evenly split between wanting to allow and deny data collection in general; hence, a classifier that takes more context into account is necessary for effective prediction. The precision and recall values are 0 because the majority class was always to *deny* data collection, resulting in no true positives ever being predicted, which is clearly not representative of an individual’s actual preferences.

Understanding how well we can predict an individual’s decision to allow or deny data collection is useful in applications such as where a system pre-populates a privacy control panel with an individual’s predicted responses. If an individual changes a pre-populated control (i.e., responding with something different than the system’s prediction), the system can update its model with this new “correct” answer. Iteratively refining answers until the system is very confident about a decision will ultimately lead—our results suggest—to the majority of answers specific to an individual being predicted with high confidence.

3.2.4 Data Collection Notification Preferences

We presented participants with questions asking how often they want to be notified about a data collection with three different frequencies. The frequencies are whether they would want to be notified 1) every time, 2) once in a while, or 3) only the first time the data is collected. They were asked to answer their preferences for all three types of notifications on a five point Likert scale ranging from “Strongly Agree” to “Strongly Disagree.”

The best models for describing the three frequencies of notifications revealed that participants’ preferences for notification changes based on the factors and levels of factors. The three significant factors that were common between all the models were: `data_type`, `location`, and the interaction of these two factors. In these models positive coefficients (estimate) show likeliness of participants’ desire to get notification about a data collection.

In the free text answers, 41% of all participants mentioned that being informed would help them feel comfortable, indicated by phrases like “I would want to know...” or “If they would tell me...”. Within that group, `purpose`, a factor heavily dependent on `data_type` and `location`, was mentioned by the majority (66%) as something that they would want to be informed about. It was followed by `retention` (35%), a factor not found in the model. 15% also explicitly requested information on who would be collecting the data (code “collector”). In addition, 13% of this group wanted to be informed about who is accessing the data and 5% want to be informed about steps taken to ensure the security of the collected data. Eight percent of the participants showed some kind of mistrust related to the `purpose` of data collection described in the scenarios. This was expressed in various ways, from demanding to know “exactly” what was stored and requesting “guarantees” to asking for honesty or expressing general concern about their privacy.

P928: I like honesty, and with companies being honest and open about why they are sharing data, it makes it a lot easier for me to be comfortable.

More detailed information was also requested about potential risks and how their data was protected against misuse.

Notification Every Time

We measured participants’ preferences to get notified about a type of data collection every time it occurred by their answers to the question “I would want my mobile phone to notify me every time this data collection occurs.” The factors in the order of their size of effect are shown in Table 3.8. The most effective factor in explaining participants’ desire to be notified every time was the interaction between `data_type` and `user_perceived_benefit`, while the factor that had the smallest effect size was `shared`. Looking at the levels of these factors, it seems that participants were most likely to want to be notified every time when their biometrics were being collected for an unspecified purpose. Also, knowing that the data was being shared had the least effect on participants’ desire to be notified every time the data collection occurred.

Factor	Estimate	Std Err	Z-value	p-value	BIC
<i>data:user perceived benefit</i>					13467
<i>baseline=friend's house:not beneficial</i>					
biometric:not beneficial	0.09	0.21	0.46	0.64	
presence:not beneficial	-0.49	0.24	-2.04	0.00	
temperature:not beneficial	-0.38	0.35	-1.10	0.27	
video:not beneficial	0.48	0.22	2.19	0.00	
biometric:purpose unspecified	0.88	0.42	2.12	0.01	
presence:purpose unspecified	-0.04	0.48	-0.08	0.93	
temperature:purpose unspecified	-0.71	0.46	-1.55	0.12	
video:purpose unspecified	-0.19	0.47	-0.42	0.67	
<i>data:happening within 2 years</i>					13591
<i>baseline = friend's house:disagree</i>					
video:agree	-0.48	0.34	-1.44	0.15	
biometric:agree	-0.01	0.24	-0.04	0.96	
presence:agree	-0.76	0.33	-2.31	0.02	
temperature:agree	-0.11	0.39	-2.28	0.78	
<i>being shared:data (baseline = not being shared:specific position)</i>					13738
<i>being shared:data</i>					13738
<i>baseline = not being shared:specific position</i>					
being shared:presence	0.96	0.22	4.39	0.00	
being shared:temperature	-0.27	0.20	-1.32	0.18	
being shared:video	0.73	0.17	4.20	0.01	
<i>data (baseline = specific position)</i>					14198
biometric	0.17	0.44	0.39	0.70	
presence	-0.57	0.54	-1.07	0.29	
temperature	-1.66	0.54	-3.07	0.00	
video	-0.02	0.52	-0.03	0.98	
<i>happening within 2 years (baseline = disagree)</i>					14697
agree	-0.27	0.19	-1.42	0.15	
<i>user perceived benefit (baseline = beneficial)</i>					14923
not beneficial	0.89	0.16	5.45	0.00	
purpose unspecified	0.69	0.35	1.94	0.04	
<i>benefit of scenario:location</i>					15281
<i>baseline = benefit to company:friend's house</i>					
benefit to user:department store	-0.01	0.25	-0.02	0.98	
purpose unspecified:department store	0.13	0.28	0.46	0.65	
benefit to user:house	-0.65	0.27	-2.38	0.01	
purpose unspecified:library	0.71	0.22	3.18	0.00	
benefit to user:library	0.31	0.25	1.28	0.20	
benefit to user:public restroom	0.16	0.25	0.62	0.54	
benefit to user:work	0.29	0.24	1.18	0.23	
<i>benefit of scenario (baseline = benefit to company)</i>					15421
benefit to user	-0.26	0.41	-0.66	0.51	
purpose unspecified	-0.77	0.36	-2.12	0.00	
<i>location (baseline = friend's house)</i>					15471
library	-1.11	0.19	-5.58	0.01	
work	-1.09	0.19	-5.57	0.00	
house	0.79	0.21	3.81	0.00	
department store	-0.69	0.20	-3.41	0.03	
public restroom	-0.29	0.19	1.48	0.14	
<i>being shared (baseline = not being shared)</i>					15539
being shared	0.17	0.11	1.62	0.11	

Table 3.8: Generalized Linear Mixed Model Regression output for every-time notification. A positive coefficient (estimate) shows likelihood of participants' desire to get notification about a data collection every time. Factors are ordered by their contribution: the factor with the lowest BIC contributes most to explain participants' preferences about every-time notification.

Notification Once in a While

We measured participants' preferences to being notified only once in a while about a type of data collection by their answers to the question "I would want my mobile phone to notify me

Factor	Estimate	Std Err	Z-value	p-value	BIC
<i>data (baseline = specific position)</i>					14172
biometric	-0.56	0.16	-3.35	0.00	
presence	-0.07	0.24	-0.27	0.78	
temperature	-0.03	0.25	-0.13	0.90	
video	-0.42	0.14	-3.07	0.01	
<i>IUIPC</i>					
control	-0.29	0.07	-4.03	0.00	14231
<i>location (baseline = friend's house)</i>					14238
library	0.48	0.22	2.21	0.02	
work	0.64	0.18	3.63	0.00	
house	0.31	0.19	1.63	0.10	
department store	0.29	0.22	1.36	0.18	
public restroom	0.26	0.22	1.19	0.23	
<i>data:location</i>					14243
<i>baseline=specific position;friend's house</i>					
biometric:department store	0.24	0.21	1.14	0.26	
biometric:library	-0.02	0.20	-0.09	0.92	
presence:department store	-0.62	0.29	-2.14	0.00	
presence:home	-0.001	0.27	-0.01	0.99	
presence:library	-0.85	0.29	-2.83	0.00	
presence:public restroom	-0.67	0.29	-2.29	0.03	
presence:work	-0.48	0.25	-1.87	0.61	
temperature:department store	-0.76	0.38	-1.98	0.00	
temperature:home	0.52	0.28	1.86	0.62	
temperature:library	-1.34	0.33	-4.06	0.00	
temperature:public restroom	-0.86	0.31	-2.87	0.00	
temperature:work	-0.87	0.28	-3.12	0.04	
video:department store	-0.09	0.19	-0.48	0.62	
video:library	-0.11	0.19	-0.54	0.59	
video:public restroom	-0.30	0.25	-1.20	0.22	

Table 3.9: Generalized Linear Mixed Model Regression output for once-in-a-while notification. A positive coefficient (estimate) shows likelihood of participants’ desire to get notification about a data collection every once in a while. Factors are ordered by their contribution: the factor with the lowest BIC contributes most to explain participants’ preferences for once-in-a-while notification.

every once in a while when this data collection occurs.” The results in the order of effect size are shown in Table 3.9. The model selection algorithm showed that the most effective factor in explaining participants’ desire to be notified once in a while was `data_type` and the least effective factor was the interaction between `data_type` and `location`. The coefficients of the levels within these factors show that participants were most likely to want to be notified every once in a while when their biometric was being collected and their desire to get notification every once in a while was least effected by knowing that their presence was being collected while they were at a department store.

Notification the First Time

We measured participants’ preferences to being notified only the first time about a type of data collection by their answers to the question, “I would want my mobile phone to notify me only the first time this data collection occurs.” Table 3.10 shows the factors we got from the model selection in order of the effect size. The most effective factor in explaining participants’ desire to be notified for the first time was `user_perceived_benefit` and the factor with the smallest effect size was the interaction between the `data_type` and `location`. More specifically,

Factor	Estimate	Std Err	Z-value	p-value	BIC
<i>user perceived benefit (baseline=beneficial)</i>					14487
not beneficial	-0.47	0.07	-7.09	0.01	
purpose unspecified	-0.32	0.05	-6.08	0.00	
<i>location (baseline=friend's house)</i>					14567
library	0.74	0.22	3.37	0.02	
work	0.86	0.18	4.76	0.00	
house	0.08	0.19	0.41	0.68	
department store	0.75	0.22	3.36	0.03	
public restroom	0.61	0.22	2.81	0.00	
<i>data (baseline=specific position)</i>					14587
biometric	0.17	0.17	1.02	0.31	
presence	0.78	0.24	3.24	0.00	
temperature	0.81	0.25	3.30	0.00	
video	0.00	0.13	-0.02	0.99	
<i>data:location</i>					14617
<i>baseline = specific position:friend's house</i>					
biometric:department store	-0.58	0.21	-2.79	0.00	
biometric:library	-0.30	0.20	-1.51	0.13	
presence:department store	-1.05	0.29	-3.66	0.00	
presence:home	-0.23	0.27	-0.83	0.41	
presence:library	-1.19	0.29	-4.02	0.02	
presence:public restroom	-1.19	0.29	-4.13	0.00	
presence:work	-0.48	0.25	-1.86	0.06	
temperature:department store	-1.61	0.38	-4.26	0.00	
temperature:home	0.23	0.28	0.82	0.41	
temperature:library	-1.35	0.32	-4.18	0.00	
temperature:public restroom	-1.09	0.31	-3.58	0.00	
temperature:work	-1.17	0.28	-4.19	0.01	
video:department store	-0.16	0.19	-0.85	0.39	
video:library	-0.17	0.19	-0.89	0.37	
video:public restroom	-0.54	0.25	-1.20	0.22	

Table 3.10: Generalized Linear Mixed Model Regression output for first-time-only notification. A positive coefficient (estimate) shows likeliness of participants' desire to get notification about a data collection only the first time. Factors are ordered by their contribution: the factor with the lowest BIC contributes most to explain participants' preferences for first-time-only notification.

participants were most likely to want to get a notification only the first time if the data collection was not beneficial to them. Also their desire to get notified only for the first time was least effected when their biometric was being collected while they were at a department store.

Summary of Data Collection

At the end of each survey, we asked participants the question “Keeping in mind the 14 scenarios, how often would you be interested in seeing a summary of all such data collection?” Participants could select either every day, every month, every year, or never. Answers varied, with 23% (n = 232) saying they would like a daily summary and 63% (633) selecting a monthly summary. Additionally, 8% (85) would have liked a summary every year and 6% (57) never wanted to receive one.

3.3 Discussion

Our results demonstrate varied privacy concerns, both across IoT scenarios and across participants. Our results also indicate that participants are more comfortable about data collection when classical privacy and data protection rules, such as the Fair Information Practices, are applied and individuals are given an explanation about why their data is being collected. However, other results underline the need for technology to support the awareness of data collection and that can meet the different desires for being notified.

3.3.1 Privacy Preferences Are Complex

How individuals feel about different data collection scenarios depends on various things. Individual preference play as much a role as social norms and expectations.

On one hand, our analyses show that participants are largely in agreement on a number of practices where social norms are in place that define what is acceptable and what is not. For example, participants expressed more comfort with data collection in public spaces, but rejected scenarios that described video cameras used in private rooms and shared with law enforcement. This is likely related to a long, western tradition of public/private dichotomy. However, this dichotomy is challenged by smart-home technology with centralized, cloud-based services that do not follow expectation of “what happens at home stays at home.” For example, Samsung received criticism for advising the public not to have private conversations in front of their smart TV [152] as it uses a third party speech-to-text service for voice commands. Smart-home device manufacturers should be aware and respectful of individuals’ mental models of data collection within the home and do their best to communicate practices that may be surprising to their customers.

On the other hand, we saw a large number of scenarios in which there was no clear indication of what is generally acceptable. For example, participants showed a high variance in the level of comfort with respect to the collection and storage of movement patterns at their workplace for the purpose of optimizing heating and cooling. Social norms have yet to emerge with respect to technology that has just recently become available. However, scenarios like these also reflect how individual preferences might differ in the long run. Individuals have to weigh their potential loss of privacy, due to camera surveillance against the benefit of reduced energy consumption. The complexity of this individual decision process is also reflected by the fact that our models describing the comfort level and the choice to allow or deny a data collection do not completely overlap. Here individual concerns about what might happen to the data, in combination with personal experience (e.g., how much one trusts her employer), play a role in determining whether or not one feels comfortable with the data collection and will allow it.

3.3.2 Addressing Privacy Concerns

Both the qualitative and quantitative data show that participants prefer anonymous data collection. Temperature and presence sensors produce data that are not immediately identifying and participants consistently expressed higher comfort with these scenarios. This finding was further reinforced by our free-text results, as anonymous data was the second most mentioned preference for data collection. This is further confirmed through interviews done in a previous study [46].

The relatively high discomfort with data inference, combined with high comfort regarding collection of anonymous data indicates that people may be generally unaware that with the Internet of Things it will be easier to re-identify individuals from otherwise anonymous data. In light of our findings, it is likely that this is something that would cause discomfort. This gap in understanding should be kept in mind when providing privacy information for IoT data collection.

We found that participants favor short retention times and are more comfortable when data is deleted after its purpose is met, or not kept longer than a week. Insights from the free-text responses indicate that this is related to an increased awareness of data breaches, the fear of misuse of data, and concerns regarding bad data security practices at companies. As previous research has shown, a growing number of people have already experienced misuse of their data [282]. With the growing number of IoT devices, the probability of data breaches further increases, resulting in higher concern and less trust in the technology. To address these types of concerns, IoT device manufacturers should take precautions, both technical and administrative, to protect their customers' data and communicate these practices to the public.

3.3.3 Towards Awareness and Control

Approaches for eliciting consent or providing information are less likely to work in the IoT setting. For example, a classic privacy policy cannot be shown on many types of IoT devices, such as a smart watch. Still, people demand information about the entity collecting data, the purpose of the collection, the benefit they receive from it, and the retention period of the collected data.

In open-ended responses, participants explicitly asked for transparency in data collection and its handling. Discomfort increases when data is shared with third parties or used to infer additional information. Participants want to be informed not only about the purpose of data collection and the handling of data, but also possible security risks associated. This finding is also confirmed by previous work which found through interviews that transparency about the data collected and the purpose of the collection influence comfort levels for data collection by IoT devices [46].

Additionally, our results show that how often and about what participants want to be informed is greatly dependent on individual comfort levels. But information requests also heavily depend on whether or not individuals think a use of their data is beneficial to them or serves a greater good. To answer this question even semi-automatically requires more specific and neutral information about the purpose of a data collection. We also saw that two thirds of participants would appreciate a monthly summary about what data has been collected about them (see section 3.2.4).

To develop technical support for this is a major challenge in a fractured IoT landscape that still lacks standardization. One option to streamline these efforts, at least on a smaller scale like in smart homes, would be to build upon the Manufacture Usage Description (MUD) Specification [111] to include information on purposes of data collection and simplify the aggregation of information about data collection.

Our analysis suggests that many people want to retain control of their personal data. Future IoT services should take this into consideration when designing privacy notices instead of creating more “one-size fits all” policies.

More specifically, we suggest the adoption of the idea of personalized privacy assistants (PPA) already used in the context of mobile apps [214]. A PPA may be a tool or agent running on behalf of each individual that can proactively predict their decision to allow or deny data collection, relieving the individual of making decisions when they can be predicted with high accuracy. This predictive model could be used to, i.e., pre-populate a privacy control panel with individuals' preferences. In a deployed system, we could use a form of online machine learning to continue to update the model to a specific individual's preferences. Our predictive model 3.2.3 showed that with a few data points per individual (three), we could predict the rest of their eleven answers with an average accuracy of 88%. In a deployed system, we expect the model would have more specific data points about individuals on which to base predictions, which would be even more accurate.

3.4 Conclusion

In this chapter, we took the first step toward informing users' privacy-related decision making, which was to understand what users are concerned about and what they would like to be notified about related to an IoT data collection and use scenario. We reported on a large-scale vignette study on privacy concerns and provided statistical evidence showing that participants want to be informed about certain details of IoT data collection and use, such as what types of data are being collected, what it is used for, and how long it will be stored. We asked 1,007 participants to rate realistic scenarios about data collection occurring in multiple contexts. Our results enhance the findings of previous, mostly qualitative, research with statistical evidence that identifies specific factors related to IoT data collection and use scenarios that impact individuals' privacy concerns. Among these factors are the type of data that is collected, retention time, third-party sharing, perceived benefit, and the location at which an IoT device collects data. The statistical results are confirmed by analyzing the free-text responses, which emphasize concerns regarding the collection of biometric data as well as data collection occurring in private spaces.

Chapter 4

The Influence of Friends and Experts on Privacy Decision Making in IoT Scenarios

As increasingly many Internet-of-Things (IoT) devices collect personal data, users face more privacy decisions. In Chapter 3, we discussed how factors related to IoT data collection and use scenarios (e.g., data type, purpose of data collection) impact people’s concerns and preferences. In this chapter, we discuss the impact of social influence on people’s IoT-related decision making.

From early childhood, we learn that others’ opinions and judgments are frequently a reliable source of evidence about reality [105]. In many social and biological systems, individuals rely on other members’ perceptions and observations to make decisions or adapt their behaviors accordingly [45, 66, 77, 253]. As defined by Latané, such influence is the result of “the real, implied, or imagined presence or actions of other individuals” [198]. Social influence has been demonstrated to have a strong impact on people’s decision making in many domains [68] as people look at others’ behaviors and opinions to inform and improve their own judgments [12, 125].

Research demonstrates that social cues also have an effect on users’ information-sharing behaviors on social-networking sites (SNSs). Spotwood and Hancock found that SNS users’ privacy-related behaviors and decisions are influenced by explicit social cues. For example, when users are made aware that most users select a privacy-protective setting, they are more likely to choose a more private setting themselves [312].

To mitigate the challenge of privacy-related decision making in IoT settings, we sought to understand the manifestation of social influence in IoT scenarios. Researchers have shown that reliance on social influence increases as the uncertainty in individuals’ judgments and decisions rises [323]. Such uncertainty is especially prevalent in privacy decision making in the IoT world, where data collection introduces inevitable trade-offs between risks and benefits. Social cues may help users make faster, more informed decisions by presenting information about how others have decided in similar cases.

Social influence can be categorized as either *direct* or *indirect*. *Direct* social influence is

This chapter is a lightly edited version of a paper previously published as: Pardis Emami-Naeini, Martin Degeling, Lujio Bauer, Richard Chow, Lorrie Faith Cranor, Mohammad Reza Haghighat, Heather Patterson. The Influence of Friends and Experts on Privacy Decision Making in IoT Scenarios. In Proceeding of the 21st ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW), 2018 [116].

based on persuasion, whereas *indirect* social influence is a subtler psychological process, which occurs as a result of knowing aggregate information about others' actions [236].

In this chapter, we investigate the effects of indirect social influence on user decisions about whether to allow data collection by IoT devices. In our large-scale online study, we exposed Mechanical Turk (MTurk) participants to various hypothetical data-collection scenarios and asked whether they would allow or disallow data collection in each scenario. Participants were randomly assigned to one of five conditions, which varied in the source and type of social influence. In four of these conditions, participants were told what percentage of their *friends* or of *privacy experts* had assented to the data collection. For each source of influence (friends and privacy experts), we further varied the type of the influence to be *consistent* or *inconsistent* with the decisions made in that scenario by participants who were not primed in a 500-person pre-study. For example, an *inconsistent experts* scenario would include a statement that 85% of privacy experts allowed data collection, whereas the majority of our pre-study participants who were exposed to the scenario without social influence had chosen to deny data collection. In the control condition, participants were not exposed to any social influence.

Our study design enabled us to understand the impact of indirect social influence on privacy decision making. We found that in general, displaying aggregate information about the behaviors of friends and privacy experts' sways participants' privacy-related decisions. Moreover, we found that social cues help participants make their decisions significantly faster.

To better understand the variables that predict people's response to social cues when making privacy-related decisions, we studied factors emerging from the literature such as expertise [160, 175, 244], level of consensus [232], opinion difference [244], and task difficulty [36, 73, 105].

We found that participants are influenced by both privacy experts and their friends, but in different ways. When friends denied data collection, our participants were more influenced than when friends allowed data collection. On the other hand, and perhaps surprisingly, participants were influenced by privacy experts more when they allowed data collection.

We also observed that the influence of social cues could wear off or get stronger, depending on whether the cues were consistent or inconsistent with pre-study participants' opinions for several scenarios in a row. For example, 40% of participants who were shown a single inconsistent social cue would follow that cue; but only 32% of participants who had previously been exposed to one inconsistent social cue, and 29% of those who had seen two inconsistent social cues, would follow the subsequent (again inconsistent) cue. On the other hand, if the social influence is consistent over several scenarios, then participants are more likely to be affected by it in future scenarios.

In addition, the majority of our participants specified that technology expertise was the quality that would influence them the most when making privacy-related decisions (77% of participants), whereas the least selected influential quality was having a friendship history (19% of participants). This suggests that participants' decisions would be affected significantly more by advice from individuals who are known to have more expertise than by naive cues from people with whom they have a friendship history.

The rest of this chapter is organized as follows. We first describe our methodology. We then present the outcomes of our data analysis and the resulting model. Finally, we interpret our results and conclude this chapter.

4.1 Methodology

We conducted a mixed-design online study with 1000 Mechanical Turk (MTurk) participants from the United States in order to understand the impact of social influence on privacy-related decision making regarding IoT devices and identify the factors affecting this influence. In this section, we first discuss the design of the study and then describe the approaches we used to analyze our data.

4.1.1 Study Design

We conducted our study on Mechanical Turk so that we could recruit a large number of diverse participants quickly and economically. By recruiting 1000 participants we had adequate statistical power to conduct the desired tests on the data and make comparisons across 5 treatment groups. MTurk has been used frequently for experimental studies on related topics such as understanding people’s privacy concerns and biases in decision making [115, 153, 275]. Nonetheless, MTurk does introduce some biases, which we will discuss. To improve the reliability of our results, we required MTurkers’ Human Intelligence Task (HIT) rate to be above 90%.

Before launching the main study, we ran a pre-survey with 500 MTurk participants. In that survey, we presented participants with 28 hypothetical IoT data collection scenarios, each describing a location, a data collection device, the type of data being collected, how data will be used and shared, and how long data will be retained. After each scenario, we asked participants if they would allow or deny that data collection. The factors and their interactions that we selected to test in these scenarios have been shown by researchers to be among the factors that influence privacy concerns [39, 42, 188, 201, 202, 203, 210].

From the 28 pre-survey scenarios We selected three groups of scenarios for our main study, representing a range of privacy concern levels: three *allow* scenarios, where more than 80% of our pre-study participants agreed to allow data collection (without being swayed by social influence, which was not present in the pre-survey); three *deny* scenarios where fewer than 20% allowed data collection; and three *balanced* scenarios where 45% to 55% allowed data collection. In the main study, participants were exposed to these nine scenarios, which are included in Appendix B, in random order, with a series of questions after each scenario.

In our main study, we randomly assigned 1000 participants to one of four experimental conditions or the control condition, for a total of 200 participants per condition. Participants were not presented with any social cues in the control condition, whereas in the experimental conditions, we appended information about the percentage of *influencers* who allowed the data collection — described either as *friends who use this app* or *privacy experts*. For each decision in our study, we used the average opinion of the pre-study participants as a proxy for an initial opinion on that decision. To understand how opinion difference impacts privacy-related decision making regarding IoT devices, we tested the *consistency* of the social cue. A consistent social cue reflects a small opinion difference, whereas an inconsistent social cue exhibits a large opinion difference. In the two consistent conditions, participants were told that most influencers allowed data collection for *allow* scenarios and denied data collection for *deny* scenarios. Conversely, in the two inconsistent conditions, participants were told that most influencers allowed data collection for *deny* (resp., *allow*) scenarios. For two of the three *balanced* scenarios, participants in the

consistent conditions were told that most influencers allowed data collection; for one of the balanced scenarios they were told that most influencers denied data collection. In the inconsistent conditions, the influencers' decisions were reversed. The description of each scenario presented to the participants in the control condition was identical to what was shown in the experimental conditions, except that the sentence indicating how friends or experts had behaved was not present.

We used two levels of consensus when describing the percentage of influencers who allowed data collection: *weak* and *strong*. The weak consensus was described as either “more than 65%” or “fewer than 35%” and the strong consensus was described as either “more than 85%” or “fewer than 15%.” Participants in all four experimental conditions were exposed to the following combinations of scenarios with strong and weak levels of consensus. In the three *allow* scenarios, two scenarios had strong consensus and one had weak consensus. In the three *deny* scenarios, two scenarios had weak consensus and one had strong consensus. Finally, in the three *balanced* scenarios, two scenarios had strong consensus and one had weak consensus.

The following is scenario D1, with weak consensus, as shown to participants in the consistent friends condition:

“You are at a department store. This message is displayed on your smartphone: This store has a facial recognition system that takes pictures of customers’ faces automatically as they enter the store in order to identify returning customers. This method is used to keep track of your orders and make suggestions based on your purchasing habits. Your picture will never be deleted. Fewer than 35% of your friends who use this app allowed this data collection.”

After the participants read each scenario, we asked them to move to the next page of the survey, where we asked them six questions, as shown in Appendix B. They could return to the scenario by clicking on the back button. The first question was an attention check question designed to check whether participants understood basic information about the scenario they just read. For each participant, over the course of the nine scenarios, we asked three attention check questions about the type of data, three about retention time, and three about the location of the data collection. Following the attention check question, we asked participants whether they would allow or deny the data collection described in the scenario (the possible answer choices were allow, probably allow, deny, and probably deny) and the reasons behind their decision in a multiple choice question with 15 answer choices. In addition, we asked them on a five-point Likert scale to what extent they agree that the described data collection is beneficial to them and to what extent they agree it is beneficial to the society. Finally, we asked participants how confident they were about their decision to allow or deny the data collection.

After participants had seen all nine scenarios and answered the questions about each, we asked participants to self-report how much they were influenced by the decisions that the influencers made in the scenarios and also asked about the reasons they were or were not influenced.

For the last (ninth) scenario, we also asked participants how their decisions might change if they were shown the scenario with the same consensus level but with a different influencer (i.e., if they had privacy experts as their influencers throughout the survey, we asked them about their friends as the influencers and vice versa). We then asked them how their decisions might change if they were shown the scenario with the original influencer, but with the opposite majority decision (from more than 85% to fewer than 15% and vice versa or from more than 65% to fewer than 35% and vice versa).

We expected there could be more groups or individuals besides those we asked about in the survey questions. Hence, as an open-ended question, we asked the participants to name other potential influencers when making a privacy-related decision.

As trust is known to play a role in response to social influence [165], we asked participants to specify their level of trust in a number of potential influencers, such as privacy experts, their family, or their colleagues. The list of potential influencers was derived from a pre-survey question in which participants were asked in an open-ended question to describe people or organizations they would be interested in consulting to help them make a similar decision.

Next, we asked participants “What qualities would make you likely to be influenced by a specific group of people when you need to make decisions like the ones in our scenarios?” We provided a list of nine qualities and invited participants to specify their own.

At the end of the survey, we asked participants some general demographic questions about their age, gender, education level, and income range.

In the control condition, the questions we asked at the end of each scenario were identical to what we asked in the experimental conditions. Only in the ninth scenario did we ask participants to assume having privacy experts and their friends as influencers, and we again posed the same questions that we asked the participants in the experimental conditions.

To have a record of how much time participants spent throughout the study, we instrumented our surveys to collect the time by setting invisible timers before each question.

4.1.2 Data Analysis

One of our main goals in this study was to find out what factors explain whether or not participants follow the social cues they receive from privacy experts and their friends. The complete list of the factors that we analyzed in our study and their corresponding levels are described in Table 4.1.

We conducted a mixed between-subjects and within-subjects study with experimental factors between participants and repeated measure factors within participants. Thus, we applied a mixed-model logistic regression with both random intercept and random slope on a binary outcome to describe whether participants acted consistently with the influence (1) or not (0). To avoid over fitting, we performed an exploratory analysis on only the first 20% of the data [162]. We looked at the distribution of factors and the summary statistics to find trends in our collected data. We also applied the model selection on the first 20% of the data. In order to find the best model, we performed backward elimination and compared the models by their Bayesian information criterion (BIC), which is a general metric for the goodness of fit. We took the following steps to select the best model [176]:

1. Start by building the model with all the factors and interaction terms.
2. Remove the factor or the interaction of two factors which has the highest p -value. If the interaction of a factor that has the highest p -value is still in the model, it will not be removed until all its interactions are removed from the model.
3. Repeat step (2) until the BIC does not decrease.

After finding the best model, we checked the performance of our model by training the model on the first 80% and testing the model on the last 20% of the collected data [94].

Factor	Levels
influencer type	privacy experts, friends
how consistent the social cue is compared to the responses of pre-survey participants	consistent, inconsistent
strength of the social cue	strong (more than 85% or fewer than 15%), weak (more than 65% or fewer than 35%)
total number of prior scenarios	0 to 8
direction of the social cue	toward allow, toward deny
to what extent the data collection is beneficial to me	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent the data collection is beneficial to society	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent participants trust [specific groups] e.g., their friends, privacy experts, or colleagues	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent participants agree that [privacy experts/friends] have more technical knowledge than they do	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent participants agree that [privacy experts/friends] have more background information than they do	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent participants agree that they generally make decisions on their own	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
current scenario type	allow, deny, balanced
prior scenario type (note: this is determined by condition for allow and deny scenarios but will vary for balanced scenarios)	consistent, inconsistent
to what extent participants agree that they have sufficient knowledge about privacy	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
to what extent participants agree that they have sufficient knowledge about technologies mentioned in the scenario	strongly agree, agree, neither agree nor disagree, disagree, strongly disagree
general demographic information such as: age, gender, income range, and education level	the corresponding levels for each demographic factor are presented in Appendix B

Table 4.1: Description of the data analysis factors.

4.1.3 Free Text Responses

As part of the survey, participants were asked to specify which other people or organizations would influence their decisions. The answers were collected as free text and later annotated by two researchers. Each annotator first independently developed a simple codebook [222] based on 250 answers, 50 from each of the five study conditions. The annotators then discussed, refined, and merged both codebooks into one, which took two iterations. The resulting codebook contained 23 codes and was used by both annotators to independently code all the responses. After finishing the coding, we decided to merge some codes (e.g., “spouses” and “parents” into the more general annotation “family”), as the number of their occurrences was very low and there was little conceptual difference between the codes. The resulting 14 codes are listed in Table 4.2. The annotator agreement as measured by Cohens Kappa was $\kappa = 0.81$, which is regarded as very good to excellent [195].

In addition to manual coding, we also examined the sentiment of the answers using an online service¹ that classifies the sentiment of a given text as *positive*, *negative*, or *neutral* to learn whether participants expressed any strong feelings toward the question.

4.1.4 Limitations

We conducted our study using the Mechanical Turk platform. Although the demographic information of our sample of MTurkers was close to the average U.S. population, our sample was not representative of the U.S. population. For instance, MTurkers are both younger, more educated, and more privacy-sensitive than the overall U.S. population [177, 286].

Researchers also worry that MTurkers do not devote full attention to the questions they are asked [153]. To mitigate this issue, we instrumented our surveys with attention check questions for each scenario. Upon examining participants’ responses to the attention check questions and their response times, we confirmed the success of our approach. Despite all the limitations of the MTurk population, prior work has confirmed the reliability of the responses [59]. In addition, research has shown that the MTurk population exhibits the same decision-making biases as the general population [153].

Another limitation of this study was that we asked participants to imagine themselves in nine hypothetical data collection scenarios followed by social cues. The main reason that we applied the vignette-based methodology was to control the factors that we were interested in studying. We acknowledge that the context of the vignettes was not as detailed or realistic as real-life scenarios. However, we wanted to conduct a carefully controlled study and examine specific relevant factors in simplified data collection scenarios, whereas adding more context to the scenarios would have introduced some confounding factors that we could not control in our statistical analysis. Our study provided statistical evidence that social influence indeed plays an important role in privacy-related decision making. Now that we have demonstrated which effects exist in these scenarios, future work should explore richer and more realistic contexts.

The focus of our study was to understand the impact of social cues from friends and privacy experts. However, there are other interesting groups or individuals mentioned by our participants that are worth investigating in future studies, such as family members or colleagues.

¹We used text-processing.com that offers a sentiment classifier trained on tweets and movie reviews.

Code	Occurrences	Description
no one	265	Participants do not want any input
experts	217	Privacy experts were mentioned most often, but this group includes also other experts like safety or technology experts
family	178	Includes mentions of parents, spouses, siblings, or family in general
law enforcement	97	Mostly mentioned in a general way, but some participants referred to specific institutions like police, FBI, or NSA
media	52	Participants said they would be looking for news, some referring to specific platforms where they read online reviews
friends	47	Some participants tried to differentiate to emphasize that this group should consist of “close” or “trusted” friends
coworkers	39	Especially referring to workplace scenarios, participant would ask colleagues about their opinions
government	34	Those expected guidelines from government officials on what is appropriate
companies	23	Some wanted to know more about the reasons for a data collection, therefore referring to the companies asking for their data
non profits	22	Most often mentioned were EFF or ACLU
general public	19	Mentioned interest in what the “general public” or “society” would do in these scenarios
boss	17	Similar to “coworker,” some participants would listen to what their managers or superiors at work would recommend
celebrity	15	Specific and unspecific mentions of celebrities. Most notably Edward Snowden (6 times)
don't know	12	Participants had no preference

Table 4.2: Descriptions of the codes used on free text answers. An occurrence is counted if both annotators used the same code.

Gender		Age		Education		Income	
Male	51.1%	Range	18-74	No high school degree	0.0%	< \$25k	26.0%
Female	48.6%	Mean	35.1	High school degree	29.4%	\$25k-\$49k	34.6%
Other	0.0%	Std. Dev.	10.2	College degree	47.1%	\$50k-\$74k	24.7%
Prefer not to answer	0.3%			Professional degree (Masters/PhD/medical/law)	12.8%	\$75k-\$99k	10.4%
				Associates degree	9.5%	\$100k-\$124k	0.4%
				Prefer not to answer	0.0%	\$125k-\$149k	0.2%
						\$150k-\$174k	0.0%
						\$175k-\$199k	0.0%
						> \$200k	0.0%
						Prefer not to answer	3.7%

Table 4.3: Participant demographics.

Finally, the described data collections in our study were hypothetical, and hence did not impose any actual risk to the privacy of participants. Therefore, real-world concerns and decisions about IoT data collections may be different from reported behaviors based on perceived risks and benefits. Despite these limitations, we believe our results provide useful insights that can inform privacy assistant design.

4.2 Results

In this section, we present our findings. We first report on the impact of social cues on the response time for decision making (Section 4.2.1). Next, we describe and evaluate our model that predicts whether participants will follow social cues (Section 4.2.2). We compare participants’ self-reported perceptions of how they were influenced by social cues to their observed behavior and note interesting divergences (Section 4.2.3). We then elaborate on the extent to which participants report trusting different influencers and the characteristics of influencers that affect that trust (Section 4.2.4).

For our main study, we recruited 1000 Mechanical Turk (MTurk) participants from the United States. Participants took an average of 15 minutes to complete the survey. Participants’ demographics are shown in Table 4.3.

Out of 1000 participants, only 14 participants made more than two mistakes on the nine attention-check questions. However, these participants’ answers to other survey questions, and the amount of time they took to answer them, did not suggest inattention. Therefore, we did not exclude any of these participants from our analysis.

4.2.1 Faster Privacy-Related Decision Making

Research has shown that people are more likely to look for guidance and information from others when a task is perceived as difficult [36, 73, 105]. In our study, the task was the privacy-related decision to allow or deny the data collection in each scenario. We used mean response time (RT), the time it took participants to make the decision in each scenario, to measure the difficulty of each task. RT distributions are positively skewed, which contradicts the assumptions behind

some common statistical tests. Hence, when using RT as a dependent variable in analyses, we apply a generalized linear mixed model (GLMM) on the raw timing data [216]. To model the influential factors that impact the RT of decision making, we applied GLMM with a random intercept for each user. Our dependent variable was the amount of time participants spent on making decisions to allow or deny the data collection and the independent variables were factors such as the study condition and scenario type. In our model, we used a Gamma distribution, as it is commonly used to statistically describe RT distributions [340]. Participants spent 3.83 seconds on average to make the decision to allow or deny each data collection. Our analysis showed that participants who were in the experimental conditions spent 3.78 seconds on average per decision and were significantly faster than the participants in the control condition (mean = 4.24s, std. dev. = 19.62s, coefficient = -0.07, p -value < 0.05).

Drilling down, we observed that participants on average made faster decisions in the *allow* scenarios (mean = 3.69s, std. dev. = 12.62s) than the *deny* scenarios (mean = 3.91s, std. dev. = 16.75s). Compared to the *allow* scenarios, it took them significantly longer to make decisions in the *balanced* scenarios (mean = 4.02s, std. dev. = 18.17s, coefficient = 0.06, p -value < 0.05).

We observed that social cues resulted in faster decision making for all three types of scenarios. Our analysis specifically showed the significant impact of social cues on the *balanced* scenarios, which required more difficult decisions as they generally required participants to consider trade-offs between clear risks and clear benefits. Notably, participants made significantly faster decisions about *balanced* scenarios in the experimental conditions (mean = 3.89s, std. dev. = 10.03s) than the control condition (mean = 4.55s, std. dev. = 17.43s, coefficient = -0.09, p -value < 0.05), suggesting that social cues allowed participants to reach a decision more quickly. Summary statistics for the timing data are presented in Table 4.4.

Conditions	Scenario type					
	<i>allow</i>		<i>deny</i>		<i>balanced</i>	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
control	3.94	11.06	4.24	12.51	4.55	17.43
consistent experts	3.93	13.93	3.94	15.39	3.88	4.91
inconsistent experts	3.68	4.32	3.72	5.06	3.51	11.46
consistent friends	3.30	11.09	3.81	7.15	4.07	3.26
inconsistent friends	3.60	4.08	3.84	9.67	4.11	10.76
all conditions	3.69	12.62	3.91	16.75	4.02	18.17

Table 4.4: Summary statistics for response time (RT) of decision making.

In general, our analysis demonstrates that providing participants with social cues, either from privacy experts or their friends, will help them make privacy-related decisions faster, especially in more complex scenarios, which exhibit inherent trade-offs between risks and benefits.

4.2.2 Inferred Influence

After presenting each of the scenarios, we asked participants whether they would allow the data collection. These responses allowed us to infer the amount of influence social cues had in each experimental condition. For more statistical power, we binned the answers as 0 (merging “probably deny” and “deny”) or 1 (merging “probably allow” and “allow”).

For each scenario, we compared participants’ preferences to allow or deny the data collection in the experimental conditions with the preferences of participants in the control condition. We created a factor called *follow* that indicates whether participants’ decisions followed the presented social cue. This binary factor was either 0 (not follow) or 1 (follow). We observed that 63% of participants followed the social cues they received in the experimental conditions. Table 4.5 shows the differences between the fraction of participants who allowed data collections in each experimental condition and the fraction of those who allowed the same data collections in the control condition. To statistically analyze the extent of the influence in different conditions in our repeated measures study, we applied mixed-effects logistic regression with random intercept for each user. The dependent variable in our analysis was binary, indicating allow or deny, and the independent variables were the scenario type, study condition, and strength of the social cue. The goal of the regression was to determine which experimental conditions would significantly increase the likelihood of allowing or denying the data collection compared to the control condition. Using random intercept in these analyses enhances the credibility of the results as the method considers the correlation between multiple data points within each user. Our results showed that compared to the control condition, participants were most influenced in the *balanced* scenarios. Perhaps surprisingly, we found that strong inconsistent social cues from friends have the most influence on participants making more complex decisions.

In order to understand the factors that contribute to following social cues, we ran a regression analysis to build a model that describes the participants’ behavior. In this model, the dependent variable was *follow*. Besides the factors described in Table 4.1, we also included the participants’ demographic information in the model. Based on the results of our model selection process, we identified the factors that predict whether participants follow social cues. Our model showed that participants follow inputs from their friends and privacy experts differently based on whether the influence is in the direction of allowing or denying data collection. If the presented social cue favors allowing the data collection, participants will follow privacy experts more than their friends. On the other hand, when the direction of the social cue is toward denying the data collection, they will be more influenced by their friends. This difference between following experts and friends is also reflected in Table 4.5. Another significant factor in our model was the strength of the social cue. As expected, we found that a strong cue influences participants more than a weak cue.

The consistency of the cue was another statistically significant factor that contributed to our model. The regression results indicated that participants will follow consistent social cues significantly more compared to inconsistent social cues. In addition, we found that participants will become more influenced after experiencing a repeated sequence of cues that are consistent with pre-study participants’ decisions. On the other hand, participants will become less influenced after experiencing a sequence of social cues that are inconsistent with pre-study decisions, especially when the cues come from experts and suggest less privacy-protective decisions.

Influencers	Consistency	Level of consensus	Scenario type			
			<i>allow</i>	<i>deny</i>	<i>balanced – allow</i>	<i>balanced – deny</i>
experts	consistent	strong	0.11	*	0.18	−0.08
		weak	*	*	*	—
	inconsistent	strong	*	*	−0.17	0.1
		weak	*	0.15	−0.13	—
friends	consistent	strong	0.05	−0.07	0.09	*
		weak	*	−0.06	*	—
	inconsistent	strong	*	*	−0.23	*
		weak	*	0.08	−0.09	—

Table 4.5: Differences between the fraction of participants who allowed data collections in the experimental conditions and the control condition. Positive numbers indicate more participants allowed data collection, whereas negative numbers indicate more participants denied data collection. We applied generalized linear mixed model (GLMM) regression with random intercept for each participant on users’ preferences to allow or deny the data collection in order to find out whether different factors increase or decrease the likelihood of allowing or denying the data collection. The * signs inside the table indicate that the difference is not statistically significant. The — signs indicate scenarios not tested. All numeric values shown indicate statistically significant difference.

Among the factors that we tested during model selection, many turned out not to be statistically significant and some were removed during model selection. For example, none of the demographic factors were statistically significant. The detailed results of the logistic regression for our best model are presented in Table 4.6, along with the complete list of factors that we removed based on their contribution to the model.

To evaluate its performance, we trained our model on the first 80% of our dataset and tested on the last 20%. The model achieved a test AUC of 0.81, which is considered excellent [166].

4.2.3 Reported Influence

After participants had been exposed to all nine scenarios and answered the questions associated with each, we asked them to report on a five-point Likert scale how much their decision making was (or would have been, in the control condition) influenced by knowing privacy experts’ or their friends’ decisions. We found that the percentage of participants who reported that they would be influenced by privacy experts was similar in the control condition (56%) and the condition in which participants received consistent cues from privacy experts (52%). On the other hand, the percentage of participants who reported being influenced by consistent cues from privacy experts (54%) is significantly larger than the percentage of participants who reported being influenced in the other three experimental conditions (consistent (21%) and inconsistent cues from friends (5.5%) and inconsistent cues from privacy experts (24%)) (p -value < 0.05). Many more participants reported being influenced by consistent social cues from either friends or experts (119 participants) than inconsistent ones (49 participants) (p -value < 0.05 for differ-

Factor	Estimate	Std Err	Z-value	p-value
(Intercept)	1.50	0.16	9.36	0.00***
strong cue	0.09	0.11	0.85	0.39
social cues from friends	-0.04	0.09	-0.43	0.66
total number of prior scenarios	0.04	0.01	2.96	0.00**
direction of the cue: toward deny	0.27	0.09	3.07	0.00**
inconsistent prior scenarios	-0.82	0.11	-6.98	0.00***
current scenario type: <i>deny</i>	0.00	0.12	0.00	0.99
current scenario type: <i>balanced</i>	-0.11	0.13	-0.81	0.41
making decisions on your own	-0.29	0.10	-2.84	0.00**
inconsistent social cue	-1.39	0.07	-19.66	0.00***
strong social cue in the <i>deny</i> scenarios	-0.20	0.17	-1.21	0.22
strong social cue in the <i>balanced</i> scenarios	0.15	0.16	0.97	0.02*
friends' behavior toward denying the scenario	0.25	0.12	1.99	0.04*
increase in the number of inconsistent prior scenarios	-0.13	0.02	-5.86	0.00***
Observations	6400			
Log-Likelihood	-3046.915			
Akaike Inf. Crit.	6119.831			
Bayesian Inf. Crit.	6206.292			
Note:	* $p < 0.05$	** $p < 0.01$	*** $p < 0.001$	
Insignificant factors from Table 4.1 in model selection:	(6, 7, 8, 9, 10, 14, 15, 16)			

Table 4.6: Regression results for the model to explain *follow* (i.e., whether participants follow advice). These results are reported on the last 80% of the dataset (the first 20% having been used for model selection).

ences between consistent and inconsistent cues from friends; and from experts). The extent of participants' reported influence in each study condition is presented in Figure 4.1.

After asking participants to report how much they were influenced, we asked them to provide us with their reasons. Two top reasons for participants who reported being influenced were “I generally like to find out what other people have done when making a decision” (69%) and “I think my friends/privacy experts have more technical knowledge than me” (54%). Among the participants who reported not being influenced by the social cues, the most common reasons were “I generally make decisions on my own” (81%) and “I generally make these kinds of decisions on my own” (76%).

4.2.4 Willingness to Trust Influencers

After exposing the participants to all nine scenarios, we asked them whom they would trust to give them good advice when making privacy-related decisions regarding data collection. Participants were instructed to choose an answer on a five-point Likert scale from “strongly agree” to

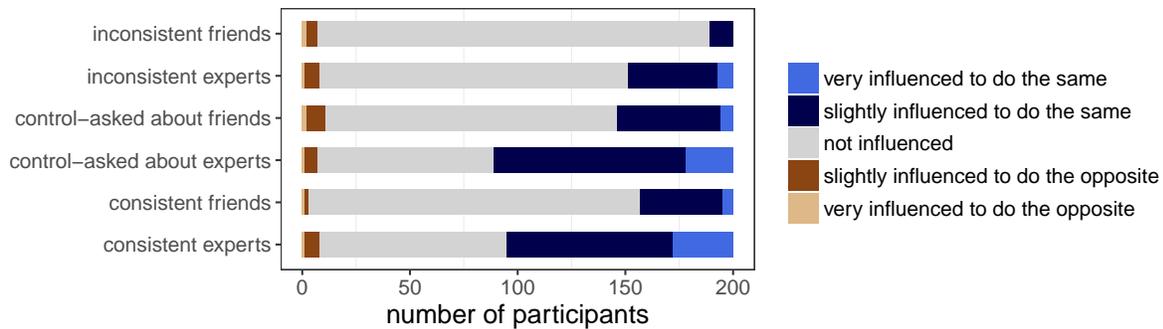


Figure 4.1: Participants’ reported influence in our five study conditions from “very influenced to do the same” to “very influenced to do the opposite.”

“strongly disagree” for each of the following groups: privacy experts, family, real-life friends, people working in technical fields, colleagues, social-network friends, and no one except themselves.

Among our five conditions, we did not observe any statistically significant difference between trusting advice from family, people working in technical fields, colleagues, and social-network friends. However, the differences between the conditions were statistically significant in trusting privacy experts, friends, and no one but myself, as discussed next.

Our results showed that most participants trusted privacy experts to give them privacy-related advice, except when presented with inconsistent social cues from privacy experts. Participants significantly lost their trust in privacy experts when their behaviors were not consistent with most participants, as determined by our pre-survey results (66% of participants in the control condition trusted privacy experts, compared to 45% of the participants who received inconsistent cues from privacy experts).

Looking at the number of prior scenarios each participant saw, we found that participants were increasingly less affected by social cues as they saw more inconsistent behaviors from the influencers, especially from privacy experts. The percentage of participants who followed inconsistent social cues from privacy experts decreased by 10% (from 41% to 31%) after seeing only one inconsistent cue. For participants who received inconsistent cues from their friends, we observed a decrease of 2% (from 39% to 37%). We observed that significantly fewer participants in the consistent-experts condition specified they trust no one except themselves in making privacy-related decisions (44%) compared to participants who received inconsistent cues from privacy experts (57%).

To better understand which qualities increased the likelihood that participants were influenced, we asked individuals to select the features they seek in their influencer. The subset of qualities that covers 98% of the responses is: some background in technology, no ulterior motive, reliability, and honesty.

As we mentioned in Section 4.1.3, we asked participants to specify other groups that would influence their privacy decisions and coded the responses; the breakdown is shown in Table 4.2. The majority of answers were only coded with one code. Although we specifically asked participants about groups other than friends and experts, most participants nevertheless mentioned that they would (or would not) be influenced by friends or experts. Other frequently mentioned

groups or organizations that were not listed in previous questions were law enforcement agencies like the police or the FBI, government officials, media reporting, and non-profit organizations.

The participants' responses also highlighted that it is important to think about what constitutes a friend or an expert. One participant wrote "I'd have to know who the 'privacy experts' and specific 'friends' were" (FI104). Another person stated "I would need more information on what defines this term. What factors are they evaluating to make this decision? How much of an 'expert' are they?" (FC50). Some participants mentioned they would not be influenced by general experts but would be influenced by domain-specific experts. Others further specified that they would be influenced only by "trusted" friends, or that it would matter whether those friends have experience with a specific technology. As one participant put it, "Friends would be the most influential because I interact with them the most and can hear directly how it affects their lives" (FI60).

A surprisingly large number of participants (178) mentioned family members as important influencers. Their responses mentioned that this includes "someone from the field," a sibling who works in the technology sector, and children who are more knowledgeable. Family members are also most trusted since, as one participant wrote, "they look out for my best interest and can be fully trusted" (EC145).

The fact that coworkers and bosses were mentioned in a number of responses is likely related to the fact that participants were presented with some scenarios that take place in a work environment. Here, for some participants, what "influencing" means seems to shift. One participant said, for example, that management would influence her decisions "because they can force you to do it as part of the job requirement" (EC7).

While the majority of participants replied in a neutral manner, many of those who said they make decisions without any influence (annotated with code "no one") used more negative wording to emphasize their autonomy with respect to questions of privacy. While there was no significant difference in sentiment between the conditions in general, the data shows differences in sentiment when comparing the responses on the annotation level across conditions. On average, all responses except those annotated with "no one" were classified as 16% negative, while the answers with the code "no one" were classified as having a negative sentiment in 62% of cases. Often, the participants with negative sentiments not only rejected the idea that they would be influenced by others in making privacy decisions, but also stressed that they would not allow any of the data collections because of what they seem to represent on a societal level. As one participant put it, "What influences me in these types of situations are the authors of books such as '1984' and 'Brave New World.' Nothing good comes of keeping a population under so much close surveillance" (FI57). Others emphasized their autonomy regarding questions of privacy with statements like "I believe in making decisions all on my own, not based on what others would do" (FC48).

4.3 Discussion

The literature shows that social cues serve as an effective approach to help people make informed privacy and security decisions [106, 150, 151]. Inspired by past work, we conducted a mixed-design study to test if and to what extent people are influenced by knowing the decisions of their

friends or privacy experts in different scenarios. Our results indicated that social cues from both privacy experts and friends influence privacy-related decision making about IoT data collections. A number of factors impact the extent of this influence. For example, we found that a stronger social cue has more influence, especially in balanced scenarios that expose participants to a trade-off between risks and benefits.

4.3.1 Privacy Experts or Friends?

This study focused on the impact of influence from friends and privacy experts. The wording that we used throughout our scenarios for friends was “friends who use this app,” without specifying whether these are friends in real life or friends on social media. Future studies are needed to investigate which groups of friends are more influential than others.

In the current study, we asked participants what qualities would make them most likely to be influenced by a specific group of people. The most frequently mentioned qualities were having a background in technology and not having an ulterior motive. As supported both by prior research and our results, people are, in general, influenced by both privacy experts and their friends, but differently by each group. We hypothesize that people believe that experts have the knowledge needed to make good privacy decisions related to IoT, but they believe their friends are less likely to have an ulterior motive. The trust people have in experts can be destroyed quickly: Our study participants lost trust in inconsistent social cues from privacy experts significantly faster than they lost trust in inconsistent cues from friends.

Although our analysis did not reveal a significant difference between the extent of the *inferred* influence of privacy experts and friends, our participants *reported* being influenced by privacy experts significantly more than by their friends. Also, participants noted that they are most influenced by the quality of having a technology background, which is usually more prevalent among privacy experts than among friends. A possible explanation for the lack of significant difference in inferred influence is a phenomenon called *expert effect*. Researchers have shown that people’s confidence in their own opinions and decisions gradually increase as they are shown social cues from a group of experts [253]. Hence, when people share similar opinions with the experts, they may become less influenced over time by their social cues.

By examining the decision response time, we observed that receiving social cues from privacy experts and friends helped participants make decisions faster. While faster decision time might improve the perceived usability of privacy-choice interfaces and privacy-assistant tools, we must pay attention to the quality and credibility of the social cues. If a decision is made more easily because a user trusted the influencer, individuals may feel betrayed if the recommendation turns out to be against the user’s best interest. As the number of daily privacy decisions increases, users may rely more on cues that can speed their decision making. Unlike review-based platforms such as Amazon, where people read other users’ comments on a product before making purchasing decisions, people are unlikely to spend much time on each decision when they need to make a large number of real-time privacy-related decisions about pervasive IoT data collection. However, we could imagine a privacy assistant that included social cues as well as links to more information about risks and benefits. Such information might be particularly useful the first time someone encounters a new type of device that is collecting their data, or when they are surprised by the recommendation of the influencers.

4.3.2 Wisdom of Crowds

”Wisdom of the crowd” refers to the phenomenon that a group of individuals is in aggregate more and better informed than most individuals [235]. Researchers have shown that when uncertain, people look to other people’s opinions for information to form their own [105, 125]. As Allen observed, individuals may go along with decisions and beliefs that are expressed by the majority because they think that a crowd’s opinion is more likely to be correct than theirs. In some situations, conformity is constructive and appropriate, while in other situations it is not [11] and can even be detrimental [217]. In our study, we found that people are indeed influenced by social cues from the crowd (i.e., friends and privacy experts). However, crowds are not always wiser than individuals. Thus, we need to be careful about the crowds whose opinions we are collecting in order to present to people not to mislead users with incorrect or incomplete information. There are different factors that could make a crowd wiser and more accurate than individuals, including expertise and diversity [315].

A better-informed crowd is likely to provide more useful information. Therefore, it is imperative for crowds to have some expertise and background in privacy and technology, which are also the qualities desired most by participants. Moreover, participants specifically mentioned that they want to know who the experts are and what their level of expertise is. Privacy assistants may leverage the expertise of the crowd in a variety of ways. One approach might be for developers to provide an option to users that allows them to choose the types of expertise and the crowds from which they want to receive social cues. For instance, as repeatedly mentioned in the open-ended responses, some users trust social cues from non-profit organizations such as EFF or ACLU. There may be other users who want to receive cues from government officials. Another approach would be to incorporate a reputation system similar to Amazon’s rating system. In addition, to participate as an expert, participants might be required to have some privacy and technology-related certifications, such as the Certified Information Privacy Technologist certification from the International Association of Privacy Professionals. Qualified experts could then be rated by other system users.

Another important factor to make a crowd wiser is diversity of opinions. Researchers from different fields have found benefit in having different viewpoints within a group [70, 82] and reducing the redundancy of perspectives [62]. Presenting participants with diverse perspectives about a particular data collection and its risks and benefits gives them a broader understanding of the situation and helps them make more informed decisions. In our study, the consistency of (friends’ or experts’) opinions was a between-subjects factor, while a more diverse platform could include both consistent and inconsistent information.

4.3.3 Social Influence in Action

Currently, there is no deployed privacy assistant platform that can give its users the ability to opt-out from various IoT data collection across many devices. Our findings provide guidance for the design of effective privacy assistants that can help users make more informed privacy decisions quickly.

In our study, social cues from privacy experts as well as friends influenced people in their decision making regarding IoT data-collection scenarios. However, we also found that this im-

fact is dependent on factors such as the behavior of the influencer, task difficulty, consistency of the social cue, strength of the cue, and self-efficacy. We found that for the scenarios in which the benefits are generally seen to outweigh the risks, people are more influenced by cues from privacy experts, whereas in the scenarios in which the risk to the privacy is dominant, providing people with their friends' cues will have more impact. We also found that people are significantly more likely to be influenced when making decisions about balanced scenarios, which present clear trade-offs between benefits and risks. Other research about social influence has had similar findings [148].

In addition, people will follow social cues significantly more when the influencer is acting consistently with the average response (87% will follow) than when they act inconsistently with the average (21% will follow). This observation is consistent with the term *confirmation bias*. In psychology, this phenomenon is defined as the tendency that people pay more attention to information confirming their own beliefs than information they disagree with [35, 260]. As shown in the literature, people cherry pick the advice from the crowd by focusing on the opinions that are consistent with their own [357]. From the wisdom-of-the-crowd point of view, this approach can be harmful as it will block people from incorporating the majority's perspective. Several explanations have been advanced for why people may resist being swayed by outside influence. For example, it has been shown that people can incorrectly believe that the average judgment in a crowd is no more informed than the average individual's. Holding this belief is significantly related to ignoring the opinions of other people [197]. Another reason why people give more weight to their initial decision is that they know the reasons for their own judgments, but not those behind judgments of the majority. Letting go of one's own judgments and changing one's opinions has been shown to be painful and cause regrets [309]. One way to alleviate this issue is to provide more detailed, yet not overwhelming, information about the decisions.

Another important factor that affects the weight people place on other's opinions is the size of the crowd [226]. In other words, people are more influenced by larger crowds. This is something that should be taken into account by the designers of privacy assistants that provide social cues.

Self-efficacy has been shown to be a reason that people ignore influence from others. Self-efficacy makes people feel that it is unnecessary to yield to others' decisions [218]. We studied this factor and found that participants who expressed a desire to make decisions on their own were significantly less influenced by experts and friends.

The impact of a social cue is context dependent. There were several participants who said they were influenced by their colleagues in the work-related scenarios or influenced by the fire department in the scenarios in which we specifically mentioned fire-hazard prevention as the benefit.

Prior studies have shown that decisions about adopting new technologies are related to trust [146], especially when facing uncertainty or risk [363]. Our results demonstrated that after being exposed to a sequence of cues which are inconsistent with the average behavior, participants lose trust in privacy experts faster than they lose trust in their friends. Many participants also expressed that, while they would listen to arguments, they want to be independent in their decision making—sometimes with strong statements against a perceived bias in cues or towards anyone who would try to make the decision *for* them. This could be partly related to the current news cycle that is dominated by headlines about micro-targeting and companies trying to use personal data to influence web users. However, it also highlights the importance of developing trust in the

influencers and the systems presenting the cues to counter potential negative perceptions.

4.4 Conclusion

As IoT devices become more widespread and people confront choices about personal data collection, the number of decisions that they need to make become overwhelming. In this chapter, we explored how social cues help people make faster, more informed decisions regarding their privacy. To understand how people are influenced by social cues from privacy experts and friends, we conducted an online user study with 1000 Mechanical Turk participants, randomly assigned to five conditions. We presented each participant with nine hypothetical data collection scenarios. In four conditions, we showed participants what percentage of experts or friends allowed the data collection. In the fifth condition, we described the data collections without any additional information. Our statistical results confirmed the impact of social cues on people making privacy decisions. We also found that the extent of this influence is dependent on factors such as the level of privacy protectiveness of influencers' decisions and the strength of the social cues.

So far, we have explored the impact of various factors on people's IoT-related decision making. In the previously-studied IoT scenarios, we asked participants to specify their desire to allow or deny the data collection. For the rest of this thesis, we will explore how consumers make IoT-related device purchase decisions and how we can effectively convey the privacy and security practices of IoT devices to inform consumers' purchase behavior.

Chapter 5

Exploring How Privacy and Security Factor into IoT Device Purchase Behavior

While sales of IoT devices are skyrocketing [144], 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 [65, 196]. However, there is little information available for consumers who wish to seek out IoT devices that are private and secure.

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 [102, 121, 124, 209, 228, 267]. 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 [123], fuel economy and environment label for cars [119], European Union (EU) energy label for office appliances [335], Power Content Label (PCL) for electricity [64], EnergyGuide label for home appliances [137], and Lighting Facts label for light bulbs [136]. Researchers have found that standardized labels are a promising approach for informing consumers about privacy: privacy “nutrition labels” on websites [181], privacy meters in search engines [63], and a “privacy facts” checklist in an app store [183] have been shown to impact study participant decision making. However, labels proposed in the prior work were not designed for IoT devices. In addition, those labels focused solely on privacy and did not consider security factors.

To design an informative and usable privacy and security label for IoT devices, we conducted user and experts studies. This chapter describes the first step toward designing informative privacy labels, by first exploring consumers’ IoT-related purchase process.

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 example IoT devices we tested in our study and

This chapter is a lightly edited version of a paper previously published as: Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, Lorrie Faith Cranor. Exploring How Privacy and Security Factor into IoT Device Purchase Behavior. In Proceeding of the 37th ACM Conference on Human Factors in Computing Systems (CHI), 2019 [117].

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 another 200-participant Mechanical Turk (MTurk) survey to probe the influence of privacy and security information when making IoT purchase decisions.

We also 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 (e.g., a smart camera capturing images).

Almost all interviewees found our prototype labels easy to understand and capable of providing 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 more useful, are important and timely contributions as policy makers debate new IoT privacy and security regulations.

We make the following contributions in this chapter:

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

5.1 Methodology

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

5.1.1 Semi-Structured Interview Study

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

Recruitment, Selection, and compensation

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 also 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, 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 in the literature [115, 133, 204, 205, 322]. 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 variants of the label. In one label, we tuned the privacy and security information so as 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 time was forever, level of detail was identifiable). For the third label, we tried to convey a trade-off. Figure 5.1 shows a label for a hypothetical security camera with poor privacy and security practices.

Eight participants saw three variants of labels for a security camera, eight participants saw three variants for a toothbrush, and eight participants saw three variants 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 and tell us what they believe it conveyed. We asked them to circle the parts that 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.

5.1.2 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, with choices ranging 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 them various demographic questions. It took

Privacy & Security Facts

Security Camera S200
Smart++, incorporated in United States 2017
Firmware version 3.1.6 (updated June 12, 2018)

CR Consumer Reports
Overall score out of 100

55



PRIVACY

Collected data: Video, device configuration, login info

Purpose: Security, maintenance, advertisement

Retention time: Forever

Shared with: Manufacturer

Choices: None

Independent Privacy Lab Rating: ★☆☆☆☆

Level of detail for the data that is being used: Identifiable

Level of detail for the data that is being collected: Identifiable

SECURITY

Automatic updates: No

Updates lifetime: Until January 1, 2020

Choices: Configurable updates, purchase extended updates

Encrypted communication: Yes

Authentication method: Fingerprint

Internet connectivity: Required

Independent IT Security Institute Rating: ★★☆☆☆

MORE INFORMATION

i Tip(s): Register your device to receive updates

Scan QR code for manufacturer's privacy and security information



Figure 5.1: Prototype label for a hypothetical security camera with poor privacy and security practices.

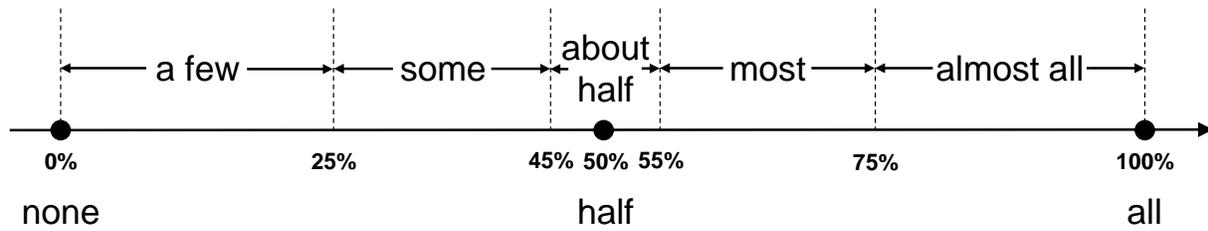


Figure 5.2: The terminology we use to report percentage of participants in Section 5.2.

participants five minutes on average to complete the survey. We compensated each participant with one dollar.

5.1.3 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 is particularly useful for semi-structured interview studies [223, 291]. 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 [131]. 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 5.2. However, to provide readers with some sense of frequency, we adopt a consistent terminology, illustrated in Figure 5.2, 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 [266]. 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. To analyze, 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.

5.1.4 Limitations

Participants in interview studies are prone to potential biases [15]. In our semi-structured interview, we avoided asking any leading questions, 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 [16, 184, 347]. 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.

5.2 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.

5.2.1 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 5.1.

5.2.2 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

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

Table 5.1: Participant demographics and IoT devices. IPA stands for Intelligent Personal Assistant (e.g., Amazon Echo, Google Home, Apple HomePod).

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.

5.2.3 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. Most of those who reported security concerns, however, had technical backgrounds and described mitigation steps they took, such as connecting their IoT devices to a

Purchase Behavior	Awareness	Knowledge	Evaluation	Concern	Management	Participant ID
Wise Proactive Protective	✓	✓	✓	✓	✓	P2, P3, P5, P8, P20, P23
Cautious Proactive Protective	✓	✗	✓	✓	✓	P15, P19
Wise Passive Protective	✓	✓	✗	✓	✓	P1, P10, P24
Cautious Passive Protective	✓	✗	✗	✓	✓	P11, P13, P17
Wise Passive	✓	✓	✗	✓	✗	P14, P16, P18, P22
Cautious Passive	✓	✗	✗	✓	✗	P12
Unconcerned	✗	✗	✗	✗	✗	P4, P6, P7, P9, P21

Table 5.2: Seven purchase behavior categories and the participants whose behaviors are described by each.

router separate from the rest of their home network.

5.2.4 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 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, when defining privacy, participants mentioned that *they* should have control over their data. On the contrary, they were mostly passive when defining security as the *device* getting hacked, except for participants with technical background, who were more proactive toward mitigating their security concerns. 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 better able to differentiate between the two. This suggests that a lack of privacy or security concerns might be attributed to not having correct and distinct definitions for these two concepts.

5.2.5 Purchase Behavior Categories

Our interview questions probed five factors related to purchase behavior: risk awareness, knowledge of privacy and security, pre-purchase evaluation of privacy and security, post-purchase concern, and post-purchase concern management. We classified interviewees into seven categories based on their responses to these questions, as shown in Table 5.2.

Some interviewees considered privacy or security in their comparison shopping, continued to be concerned about the 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 ... 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.

5.2.6 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 were 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 [113, 325]. 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.

5.2.7 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. ... 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 as 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 package of the device so that they could refer to it later. The other half 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.

5.2.8 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 5.3, 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 -value < 0.05) than for a smart thermostat. For the toothbrush, privacy and security information were significantly less influential (p -value < 0.05) compared to the other two devices. The summary statistics and the regression results for the types of devices are presented in Table 5.3. The differences we found may be due to differences in participants’ perceptions about the sensitivity of collected information.

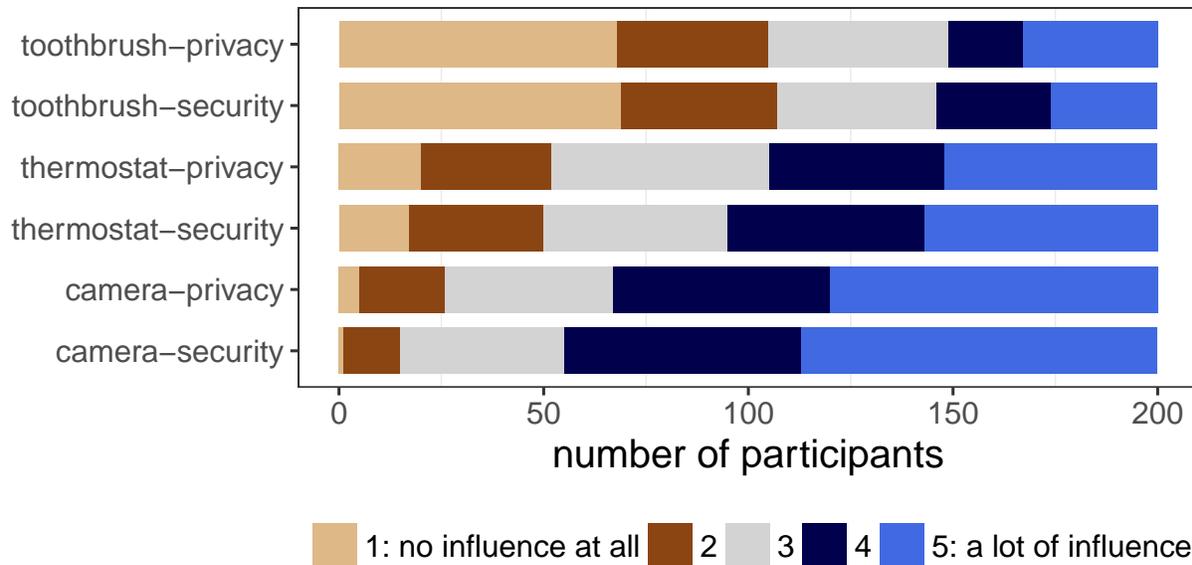


Figure 5.3: Survey participants’ responses when asked: Imagine you are deciding between two or more [IoT devices] to purchase. How much influence do you think each of the following factors would have on your purchase decision?

5.3 Discussion

We discuss ways labels can surface latent privacy and security concerns and help consumers consider privacy and security in their purchase decisions and device use. We conclude this section by discussing several design considerations for more effective privacy and security labels.

5.3.1 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 pur-

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

Table 5.3: 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 -value < 0.05).

chased. 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.

5.3.2 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, where 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 [168, 352]. Yet, 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. 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.

5.4 Conclusion

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 others 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 particularly sensitive information. Finally, we developed a prototype IoT device privacy and security label. Our interviewees found the design to be understandable. We discuss design considerations for IoT security and privacy labels and paths to adoption and enforcement.

This study showed clear interest from consumers in security and privacy labels, but also revealed consumers' lack of knowledge about what security and privacy information is actually important. In the next chapter we present a study in which we obtained input from experts to inform the content of the label.

Chapter 6

Ask the Experts: What Should Be on an IoT Privacy and Security Label?

In Chapter 5, we conducted interviews and surveys with IoT consumers to explore the importance of privacy and security in their IoT-related purchase process. Among other findings, we found that participants acknowledge the importance of considering privacy and security when purchasing IoT devices. However, not having enough information at the time of purchase prevents consumers from considering privacy and security attributes in their purchase decision making. Some resources, such as the Mozilla “Privacy Not Included” website [254] and a report published by the UK Information Commissioner’s Office [167], provide information about specific IoT devices or address limited privacy and security factors.

Critical privacy and security information could be provided to consumers by including it prominently on a privacy and security label accompanying the device. This could also increase consumers’ trust in the device manufacturer [345]. In a May 2019 proposal, the UK Digital Ministers declared their intention to mandate security labels for IoT devices, with the goal of notifying consumers about security aspects of these devices [104]. However, this plan only covers three security practices: using no default passwords, having a vulnerability disclosure program in place, and specifying the lifetime of security updates. Other proposals for IoT privacy and security labels fail to specify the specific information that consumers should be presented with on the label [102, 121, 124, 209, 228, 267, 279].

Given consumers’ scarce attention, presenting them with the most relevant security and privacy information in the most digestible form is crucial. To determine the most important information to include on IoT privacy and security labels, we solicited the opinions of privacy and security experts. In various fields, expert elicitation has been used effectively for research and decision making [29, 191, 297, 339, 341], particularly under uncertainty and when necessary information cannot be obtained from other sources [75, 163, 193].

We conducted interviews and surveys with 22 privacy and security experts. To get different perspectives, we recruited experts from industry, academia, government, and non-governmental

This chapter is a lightly edited version of a paper previously published as: Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, Hanan Hibshi. Ask the Experts: What Should Be on an IoT Privacy and Security Label?. In Proceeding of the 41st IEEE Symposium on Security and Privacy (S&P), 2020 [114].

organizations (NGOs). We also ensured that these experts come from different backgrounds related to IoT (software, hardware, and policy). We used the iterative Delphi methodology (explained further in Section 6.1) to develop a consensus among the experts around important factors and an understanding of their reasons for or against including each factor. Overall, we found that differences in opinions were driven less by fundamental differences in beliefs, but rather by differences in work experience and priorities. For example, some experts were more knowledgeable about specific security mechanisms, standards, or regulations, and prioritized factors related to their area of expertise or their organization’s mission. Prior research has shown that security experts might analyze the same artifacts differently depending on their background in specific security domains [164].

Most factors identified as important by experts are factors that they believe will inform consumers. Experts also identified some factors for inclusion that could inform experts only, mostly so that companies can be held accountable.

Prior studies suggest that layered labels can be effective [79, 117, 293, 320]. A layered label includes a *primary* layer that presents the most important and glanceable content, followed by a *secondary* layer for additional information. In our study, we asked experts to specify the layer on which the information should be included on the label. They mostly recommend putting only information that would be understandable and important to most consumers on the primary layer.

We designed a prototype layered label based on our expert elicitation study. We then conducted semi-structured interviews with 15 consumers of IoT devices (smart home devices or wearables) and presented our prototype to them. We show that all of our participants had a clear understanding of the information presented on the primary layer of the label. Although some of the factors on the secondary layer of the label were less understandable to participants who lacked privacy and security expertise, all of our participants reported that they still want such information to be included on the label mainly to be as informed as possible. In addition, our participants reported that having all the important privacy and security factors, even unfamiliar ones, on the label would help them easily search online to find more information.

We make the following contributions in this chapter:

- We distill an extensive list of privacy and security factors to identify the most important pieces of information to include on IoT labels.
- We partition the most important factors into two layers: the primary factors we want consumers to notice and consume at a glance, and the secondary factors that require more space to effectively convey risk to consumers.
- Based on our expert and consumer interviews and surveys, we propose a prototype IoT label that includes the most important factors with proposed groupings.

6.1 Methodology

We first conducted an expert elicitation study to specify the content of a privacy and security label for IoT devices. We complemented the expert study with a series of 15 semi-structured interviews with non-expert consumers and iterated on the label design. The expert and consumer interviews and surveys are provided in Appendix D.

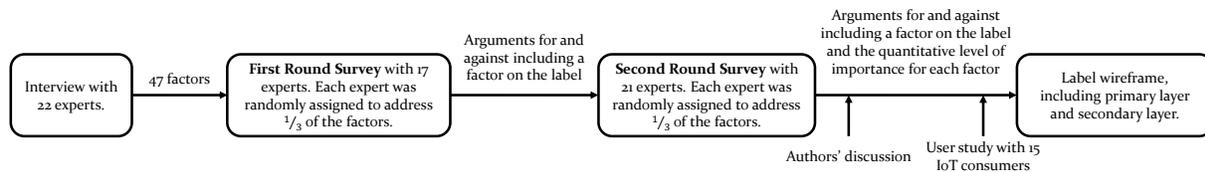


Figure 6.1: We followed a three-round Delphi method, by conducting an interview study and two rounds of surveys. Finally, we designed a label prototype that captures the findings of the process, inputs from authors’ multiple rounds of discussion, and a user study with 15 IoT consumers.

6.1.1 Expert Elicitation Study

In the expert elicitation study, our overarching goals were to identify factors that experts believed would be useful to include on a privacy and security label for IoT devices and to understand the experts’ rationale for selecting each factor. We conducted an in-depth, semi-structured interview study, followed by two rounds of surveys with 22 privacy and security experts. The process is depicted in Figure 6.1.

Participant Recruitment and Compensation

To capture a wide range of expert opinions, we recruited experts from academia, industry, government, and non-governmental organizations (NGOs) in the United States, with a diverse range of expertise: software, hardware, policy, standards, and user experience (UX). We recruited experts with whom the authors had interacted professionally or were recommended by other experts. We carefully selected experts, who are all well-known in their respective fields. More specifically, we looked for experts, who satisfied at least one of the following qualification criteria. Seven experts met two criteria.

- Computer science or engineering professor in the field of privacy and security.
- More than 10 years of research or practice in the field of privacy, security, or policy.
- Author of notable books in the field of privacy and security.
- Active involvement in cybersecurity standardization.
- Leading a corporate IoT product team.

After identifying the experts who met the qualifications we were looking for, we contacted them and invited them to either come to our institution for an in-person interview or join our interview study online over Skype. All the interviews were audio recorded and transcribed by a third-party service. We compensated experts with a \$25 Amazon gift card.

Delphi Method

As defined by Delbecq et al., the *Delphi method* is “a method for the systematic solicitation and collection of judgments on a particular topic through a set of carefully designed sequential questionnaires interspersed with summarized information and feedback of opinions derived from earlier responses” [24]. This method of qualitative research was originally developed by Dalkey and Helmer in the 1950s and has been widely used to reach consensus between a group of

experts without face-to-face interactions [90]. The Delphi method has been used in a number of studies related to policy design and implementation [5], social science [314], and human-computer interaction [245].

The Delphi method has three important features. First, the responses as well as group interactions in each round are anonymized. Second, the process involves multiple rounds of data collection procedures (e.g., interview, survey), and finally, in each round, the summary of the previous round is shown to experts as a means to reach consensus [71, 86, 87]. The study continues until consensus is reached, which generally occurs after three iterations [219].

Expert Interviews

The first phase of the Delphi method is open ended [255]. Therefore, our first step was to conduct semi-structured interviews with privacy and security experts.

We began the interviews by introducing the idea of a privacy and security label and its similarity to a food nutrition label. Following the introduction to the study and its goals, we asked experts to provide their definition of privacy and security as it relates to IoT devices. Next, we asked experts to think about the content of the label and specify the information that they think should be on a privacy and security label for IoT devices. For each piece of information they specified, we asked them to consider whether it was relevant to consumers or experts. In an iterative process, we compiled a list of security and privacy factors suggested by the experts we interviewed, and added new factors suggested during each interview. Towards the end of each interview, we presented the full list of factors so that each expert reviewed their own factors, as well as the factors suggested by previously-interviewed experts.

First Round Survey

The expert interviews resulted in an extensive list of privacy, security, and general factors that experts wanted to see on an IoT label. We then conducted a survey of the same experts to understand the rationale behind their preferences. In order to decrease fatigue, we split the factors in the survey so that each expert was presented with one-third of all the factors. For each expert, the ordering of factors was randomized.

When introducing the survey to experts, we explained that in a layered IoT label, the first or primary layer would include the most important information, and the secondary layer would contain the information on the primary layer as well as additional helpful information. We also advised experts not to worry about the design of the label when answering the questions. Then, for each factor, we asked the experts to specify whether they believe that factor is important to include on the label, and to provide reason(s) that support their answer.

Second Round Survey

For agreement over the inclusion and exclusion of factors, we conducted a second survey with the same set of experts. When introducing the second survey to the experts, we described our two key objectives for the content of the IoT label: to inform consumers, and to provide a means for holding companies accountable for their privacy and security practices. To reduce respondent

fatigue, participants answered questions for one-third of all the factors, randomly chosen from the three categories—security, privacy and general—such that they saw approximately the same number of questions within each category.

In this second survey, we used data collected from our interviews and first survey. We presented each factor from our dataset alongside the experts' reasons for inclusion or exclusion. Then, on a five-point Likert scale, we asked each expert to decide whether they believe the factor should be included on the label and to provide their rationale if different from what we presented. Next, we asked them to specify on which layer of the label they would like to place this factor and the rationale behind their choice. We asked them to classify the factor as being most relevant to privacy, security, or general information. Finally, we asked experts to provide any additional comments about the factor that came to mind.

In addition to the questions we asked for each factor, we asked experts about their opinion on separating or merging privacy and security sections on the label. At the end of the survey, we asked experts to state their privacy and security expertise and domain of knowledge, followed by some general demographic questions.

Data Analysis

We collected approximately 22 hours of interview audio recordings. We used thematic analysis to qualitatively summarize interview transcripts, following the approach suggested by Braun and Clarke [54]:

- Phase 1: A primary coder read the interview transcripts and took notes, listening to parts of the audio files as needed when the transcripts were incomplete.
- Phase 2: The primary coder created the initial codebook by examining the notes from the interview phase and the notes from Phase 1 above, listening for reasons for and against including factors on an IoT label. This step did not focus on finding patterns in the responses.
- Phase 3: The primary coder merged the smaller codes into broader themes. This step focused on finding patterns and themes from the long list of codes from Phase 2.
- Phase 4: The themes that emerged from Phase 3 were reviewed and discussed by the researchers in the group to resolve any disagreements. This step helped increase the validity of the themes. In an iterative process, some of the themes were removed from the codebook and some themes were merged into more general themes until we achieved consensus on the final themes.
- Phase 5: The finalized themes (reasons to include or exclude each factor) were moved into the final codebook.

The finalized privacy, security, and general factors were used as input to the first round of our survey, where we asked experts to provide us with their arguments. We then followed the same coding process described above to code the open-ended survey responses. After the first survey, we revised the themes (reasons for and against including a factor on the label) in the codebook and presented them to the experts in the second survey.

We reached a point of saturation in terms of finding new factors after interviewing 20 experts. In other words, no new privacy, security, or general factors were mentioned by our participants

in the rest of the interviews as well as the two follow-up surveys.

Thematic analysis is purely qualitative and inductive [54]. The literature showed that having more than one coder does not make the codes objective, since two coders could apply the same subjective perspective to the data [229]. Indeed according to a survey of CSCW and HCI publications from 2016 to 2018, only 6% of papers using thematic analysis used multiple coders and measured Inter-rater reliability [239].

An iterative, yet inductive, analysis approach in thematic analysis increases the reliability of the theme-finding process [134]. All the themes were iteratively and extensively discussed among the researchers in the group. For any disagreement, researchers traced the theme back to its corresponding subcodes and checked whether the source of disagreement was the subcodes that were used. If not, we traced the subcodes further back to experts' quotes from the transcriptions and we then decided on the appropriate subcodes and the appropriate themes arising from the subcodes. This iterative approach is recommended with qualitative methods that are high in subjectivity [134, 224]. We also improved the reliability by consulting with the expert participants using the second round survey mentioned above; this triangulation method is known as testimonial validity or member checking [291].

6.1.2 Semi-Structured Interviews with Non-Expert Consumers

We used the results of our expert study to inform the development of prototype designs for primary and secondary labels. We created prototype boxes for two fictitious brands of security cameras and included a primary-layer label on each box. We put the corresponding secondary-layer labels on a mock-up of an online shopping website. Next we conducted a semi-structured interview study with 15 non-expert consumers to gain insights into how they would use these labels and how well the labels convey risk.

Participant Recruitment and Compensation

We recruited participants by posting on Craigslist, Reddit, and our institution's recruitment website. Participants were required to be at least 18 years old and have purchased at least one smart home device or smart personal device. Prospective participants completed a short screening survey, in which we collected demographics and asked about what IoT devices they had purchased and how they purchased them. We invited a diverse sample of qualified participants to our lab for a 1-hour interview. Each interviewee received a \$25 Amazon gift card.

Initial Questions

We showed participants a box for a hypothetical security camera that did not include a label and asked them what they could tell about the privacy and security of the device by looking at its box. We then asked participants whether they had ever seen an informative label on any product. Next, we presented them with one of the two labeled security camera boxes and asked them what they could tell about the privacy and security of the device. We asked a number of questions to study participants' understanding of the content of the label and whether the information conveyed risk.



Figure 6.2: A user study participant comparing the privacy and security practices of two hypothetical smart security cameras.

Risk Communication in Comparative Purchase Process

We showed participants the other labelled security camera box, and told them this camera had the same price and features, but with different privacy and security information. We asked them to compare these two products and discuss which has better privacy and security, which device would they purchase, and the information that helped them make this decision (see Figure 6.2).

We then told participants about the secondary layer of the label, which can be accessed by scanning the QR code or typing in the URL on the first layer. After introducing the idea of the layered label, we asked participants whether they had ever seen one on any other product. We asked them to discuss the pros and cons of a one-layer and two-layer label.

Information Comprehension in Non-comparative Purchase Process

We asked participants to look at the information on the label of the product they decided not to purchase and to discuss their concerns and their understanding of the information.

Risk Communication in Non-comparative Purchase Process

We asked participants to specify the factors that seemed risky to them from a privacy and security perspective and discuss what kinds of risk they would be exposed to. We also asked how the product could be improved to reduce this risk.

Secondary-Layer Information Comprehension

We asked participants whether they would prefer to scan the QR code or type in the URL to look for additional information. Based on their preference, we scanned the QR code or typed in the URL on the primary layer to show the secondary layer to participants. We then asked them to start from the beginning of the label and tell us what each factor means to them, how useful they believe each factor would be, and if they have any suggestions to make the information more understandable.

Label Format

We asked questions about the label format, including the separation of factors into privacy, security, and general information sections. We also asked participants to specify the factors that they believed are currently misplaced, and should be either removed from the label or moved to another section or layer of the label.

Purchase Behavior

Finally, we asked questions to understand participants' purchase behavior related to online and in store shopping.

Data Analysis

We collected about 15 hours of audio recordings, which we had transcribed. The first author was the primary coder who created the codebook and kept it updated throughout the coding process. To analyze the data, we used structural coding, which is appropriate for coding semi-structured interviews [223, 291]. We defined four structural codes (e.g., attitudes toward layered labels, reasons to include or exclude a factor from the label), which we divided into 13 subcodes (e.g., being as informed as possible, lack of relevance to privacy and security). Unlike thematic analysis, structural coding is more objective, and results in a codebook used for categorization [291]. Therefore, having more than one coder and using inter-rater reliability is helpful in testing the reliability of the codebook [134]. Each interview was independently coded by two researchers, who then discussed and iteratively revised the codebook. After resolving the coding disagreements, we reached the Cohen's Kappa inter-coder agreement of 84%. Cohen's Kappa inter-coder agreement of over 75% is considered as "excellent" rate of agreement [131]. In case of disagreement, we report on the results of the primary coder.

6.1.3 Limitations

Expert elicitation is prone to overconfidence and cognitive biases [311]. To reduce overconfidence, in the second expert survey, we presented strong arguments for and against having each factor on the label so that experts could read the rationale provided by other experts before indicating their own preferences.

The experts interviewed in this study are not representative of the entire population of privacy and security experts. Our aim was to surface a wide variety of expert viewpoints. Therefore, we

recruited experts with diverse expertise related to IoT security and privacy and from different sectors. We selected experts based on our inclusion criteria, as discussed in Section 6.1.1.

In the follow-up surveys with experts, we presented each participant with only one-third of factors randomly sampled from the list of all privacy and security factors. This reduced respondent fatigue [27] and increased the quality of responses, at the cost of not being able to achieve a true consensus across all experts. In the expert study, our main objective was to collect the opinions of diverse experts, and not to reach perfect consensus. Therefore, we report themes that were only mentioned by a few experts.

We designed a label prototype based on findings from our three-round Delphi process. However, we did not conduct a fourth round of study to show experts the label and ask them for feedback. Although this would have helped us confirm experts' opinions about the factors in the context of a complete label design, it would have introduced confounding factors related to the design of the label, including, but not limited to, the order of sections on the label and the specific language used to convey the information. Since the expertise of the participants in our study was in the area of IoT security and privacy, and not in communications design, we limited the expert elicitation study to focus on the individual factors.

Our consumer study was a small-scale qualitative study designed to gain initial consumer feedback and assess the overall usefulness of the layered label approach in this context. Additional large-scale iterative design and testing is needed to refine and validate the label design.

6.2 Results

We conducted 22 one-hour, semi-structured interviews with IoT privacy and security experts with diverse backgrounds as described in Table 6.1. We compiled a list of 47 privacy, security, and general factors that experts said they would like to see on the IoT label.

We followed the expert interviews with a qualitative survey to understand the reasons experts wanted to include or exclude each factor. Out of 22 invited experts, 17 answered the first survey, with each of them being asked to comment on one-third of the 47 factors. This survey took an average of 16 minutes to complete. We collected on average seven reasons for or against including each factor on the label. We then conducted thematic analysis on the arguments provided to arrive at two or three primary reasons for and against each factor.

In the second survey, we presented experts with the reasons for and against each factor from the previous two phases and asked them to rate their enthusiasm for including the factor on either the primary or the secondary layer of the label. 21 experts participated in the second survey, spending an average of 15 minutes. We identified 12 factors that most experts recommended including on the primary layer and 13 factors most experts recommended including on the secondary layer.

Based on the expert study and authors' discussions, we designed prototype privacy and security labels for hypothetical smart security cameras and presented them to a diverse sample of 15 non-expert consumers (see Table 6.2). We asked them questions related to their understanding of the factors on the label and whether they conveyed risk. We iteratively improved the content of the label to make it more understandable, resulting in a final prototype label.

In this section, we first discuss experts' attitudes toward privacy and security. Next we present

Expert ID	Privacy & Security Expertise	IoT Focus	Workplace
P1	Privacy	Policy, standards	Enterprise
P2	Privacy	Policy, UX	Enterprise, NGO
P3	Privacy	Software	University
P4	Privacy	Policy	University
P5	Privacy	Hardware	Enterprise
P6	Privacy	Policy, software, UX	Enterprise
P7	Privacy	Policy	NGO
P8	Privacy	Policy, privacy	NGO
P9	Privacy	Policy, privacy	NGO
S1	Security	Software	University
S2	Security	Software	University
S3	Security	Policy, software	Government, University
S4	Security	Policy, security	Enterprise
S5	Security	Hardware, security	Enterprise
S6	Security	Software	Enterprise
S7	Security	Policy	Enterprise, NGO
S8	Security	Policy	NGO
S9	Security	Hardware, software	Enterprise
B1	Both	Software	University
B2	Both	Policy, software	Enterprise
B3	Both	Policy, standards	NGO
B4	Both	Policy, software	Enterprise

Table 6.1: We conducted an expert elicitation study with 22 privacy and security experts. NGO stands for non-governmental organization and UX stands for user experience.

the factors experts wanted to include on the primary and secondary layers of the label or exclude from the label, followed by a discussion of how consumers perceived those factors. We continue by discussing consumers’ attitudes toward the labels and their layered design. Finally, we present our prototype label design that we designed based on experts’ and consumers’ input.

6.2.1 Definition, Assessment, and Accountability

We started the interviews by asking experts to define privacy and security related to IoT devices. When defining security, almost all experts (21/22) mentioned the CIA triad of confidentiality, integrity, and availability. However, experts had different definitions for privacy. Some experts (9/22) defined privacy as having transparency and control over data practices and some experts defined privacy as the confidentiality aspect of security (8/22). Overall, experts’ definitions for security were mostly passive and focused on hardware and software enforcement mechanisms. On the other hand, their definitions for privacy were active and centered around policy, control, and individuals’ preferences and comfort. For instance, P7 compared privacy and security practices by saying: “If the privacy is done right, it would be more active than security because the consumer would be able to be in control.” B3 explained how privacy and security are related by

Participant ID	Gender	Age	Technical Background	IoT Devices our Participants Have Purchased
C1	F	35-44	Y	Thermostat, TV, switches, lock, outlet
C2	F	35-44	Y	Vacuum cleaner, gaming consoles
C3	M	55+	Y	Thermostat
C4	F	45-55	Y	Smart speaker
C5	M	35-44	Y	Camera
C6	M	55+	N	Smart speaker, lights
C7	M	18-24	N	Smart speaker, lights, TV
C8	M	25-34	N	Activity tracker
C9	F	25-34	N	Smart speaker
C10	F	25-34	Y	TV, camera
C11	F	18-24	N	Smartwatch, activity tracker, camera
C12	F	35-44	N	Smart speaker
C13	M	25-34	N	Smart speaker, TV, plugs, vacuum cleaner
C14	F	25-34	N	Smartwatch, activity tracker, vacuum cleaner
C15	F	25-34	Y	Smart speaker, smartwatch

Table 6.2: User study participants, demographics, and devices they have purchased.

saying: “Security mechanisms are the things that enforce the technical controls that allow us the privacy we have.”

Most experts (15/22) believed that security information is less tangible and understandable for consumers compared to privacy information, in part because it relies on technical mechanisms. S5 explained: “Consumers don’t necessarily understand some of the abstract stuff about security that they don’t see. Whereas when their privacy is breached, they are more aware of that.”

In addition, almost all experts (19/22) reported that security practices are easier to measure and assess than privacy practices, as security is more objective and less controversial, while privacy is more subjective and context dependent. P5 explained that security is easier to quantify: “Security strikes me as less subjective, and, therefore, easier to measure. Which is to say that there could be certain standards. What sort of encryption exists on the device? What encryption is in the cloud? These are all fairly quantifiable. Whereas privacy is trickier, I think. And almost ethically and morally from my point of view, there’s a lot more gray area in this.” This finding is aligned with the current efforts in IoT assessments and scoring, which are more focused toward security mechanisms than privacy practices [118, 149, 171, 322, 358].

Experts reported that IoT privacy and security labels could increase accountability (most frequently mentioned for factors on the secondary layer) and transparency. Seven experts suggested that increased transparency could be an incentive for companies to compete on privacy and security, leading to safer products. S4 explained: “There is value in forcing the company to write a list down even if the consumer doesn’t understand it. If you said, ‘list your open ports,’ there would be an incentive to make them few.”

Some experts (8/22) mentioned that IoT companies’ accountability should be different for privacy and security breaches. They said that security breaches can happen accidentally, even

if companies follow best practices. However, privacy violations could be intentional and IoT companies could even profit from them. P4 explained: “You can have the best intentions in the world, but if somebody comes up with some crazy hack overnight, you shouldn’t be held responsible for it. You should be held responsible for fixing it. As opposed to if you intentionally share someone’s data with a third party, it’s not like, oh you could have prevented it but you chose to do it.”

In the second survey, we asked experts to specify whether they prefer to see privacy and security factors in two separate sections or if we should combine them into one section. About half of the experts (10/21) believed privacy and security should be presented in separate sections. Most of these (9/10) said such separation would improve the readability and utility of the label and help educate consumers. P1 explained: “I lean toward the option of separation, because I’d like to see a streamlined label for most consumers to ‘consume’ as easily and quickly as possible.” S8 concurred: “Consumers may have preferences for one aspect more than the other and stating them separately better enables consumer choice and education.” However, the other half of the experts believed privacy and security factors should be combined into one section (11/21). For example, P4 believed that for some consumers, security seems more important than privacy. Thus, separating them on the label may cause consumers to focus only on security factors and ignore privacy information. Among those experts, who were more interested in combining privacy and security information into one section, almost all of them (9/11) mentioned that privacy and security are so related that it is not possible to completely separate these two concepts.

In the label that we presented to consumer interview participants, we grouped information into three main sections: security mechanisms, data practices, and general/more information. Participants preferred the proposed separation of sections and reported that these groupings made sense to them.

6.2.2 Factors to Include in the IoT Label

From the second survey, we found 30 factors that at least 4 out of 7 of the experts recommended including on the label (either on the primary or secondary layer) and 17 factors that at least 4 out of 7 experts recommended excluding from the label. Note that since only a third of the factors were shown to each expert (total 21 on the second survey), at least 4 responses constituted a majority.

The authors discussed the experts’ arguments and preferences for each factor and made a decision as to whether or not each factor should be included and if so, on which layer of the label. In some cases, we made a decision that contradicted the majority of experts if we felt that their arguments could be accommodated in a different way.

Primary Layer

We found 12 factors that at least 4 out of 7 experts wanted to include on the primary layer:

- Privacy rating for the device from an independent privacy assessment organization
- Security rating for the device from an independent security assessment organization

- The date until which security updates will be provided
- Type of data that is being collected
- Type of sensor(s) on the device
- Whether or not the device is getting cryptographically signed and critical automatic security updates
- Types of physical actuations (e.g., talking, blinking) the device has and in what circumstances they are activated
- Whether or not the device is using any default password
- Frequency of data sharing (e.g., continuous, on demand)
- The warranty period of the device
- Level of detail (granularity) of the data being collected, used, and shared (e.g., identifiable, aggregate)
- Access control for device and apps (e.g., none, single-user account, multi-user account)

Experts were interested in including these factors on the primary layer because they considered them necessary for consumers to know, they convey critical information about the privacy and security of the device, and they inform consumers' purchase decisions. For example, P1 explained why the type of collected data should be included on the primary layer: "I think this is the most useful information to be provided to consumers for them to compare privacy risks of IoT devices."

All of our consumer participants understood the information presented on the primary layer and were able to talk about privacy and security implications of each factor. For example, consumers associated the expiration date of the device to its security updates lifetime. C3 mentioned "planned obsolescence" when talking about the security update lifetime: "I do like the fact that you say when the security updates will no longer be available, because that alerts people to the fact that this device is going to expire. People have thermostats that last for decades and it's useful to know that this is planned obsolescence." One of our participants brought up a point of skepticism, related to how long a company claims to provide security updates: "I'm skeptical because I know that tech startups can very rarely guarantee that their servers will be online for three or more years."

Among the factors experts believed should be included on the primary layer, there were three factors that we decided to either move to the secondary layer or exclude from the label. Note that we also removed privacy and security star ratings later in the process as mentioned in Section 6.3. We expect to add them back when such assessments are available in the future.

First, we decided to move the physical actuation factor to the secondary layer because this information is not usually directly related to the privacy and security of the device. In the consumer study, all participants found this information useful from a safety point of view, but none wanted us to move this factor to the primary layer as they reported that the information conveyed by this factor does not have privacy and security implications for them.

Second, we decided to move frequency of data sharing to the secondary layer because prior work has shown that most people do not understand the privacy and security implications of the frequency of data sharing [30]. While almost all the consumers we interviewed were concerned

about data sharing, only four mentioned privacy concerns related to the frequency of sharing.

Third, we decided to exclude device warranty period from the label because it has few, if any, privacy, security, or safety implications.

Secondary Layer

We found 13 factors that at least 4 out of 7 experts wanted to include on the secondary layer of the label:

- Retention time
- Purpose of data collection
- What information can be inferred from the collected data
- Supported standards (e.g., Wi-Fi, Zigbee)
- Where the collected data is stored
- Whether or not the collected data will be linked with data obtained from other sources
- Special data handling practices for children's data
- The control that users are offered (e.g., opt-in/out from data sharing)
- Data-collection frequency (e.g., once a month, on install)
- Whether or not the device can still function when Internet connectivity is turned off
- Relevant security and privacy laws and standards to which the device complies (e.g., ISO 27001, GDPR)
- Link to the device's key management protocol
- Resource usage in terms of power and data (e.g., kw, kbps)

Experts mentioned two common reasons to include a factor in the secondary layer rather than the primary layer: the factor requires detailed information to convey risk to consumers (mentioned by 6/7) or the factor does not convey critical information related to the privacy and security of the device (mentioned by 4/7).

Among the factors our experts wanted to include on the secondary layer, there were three factors that we decided to include on the primary layer instead: date of the latest firmware update, purpose of data collection, and where the collected data is stored.

Experts wanted to have the date of firmware update on the secondary layer mainly because these updates happen frequently and the information on the label can become outdated. On the other hand, consumers need to know the firmware version to which the label is applicable. Therefore, we believe the firmware version number and date information should be provided on both layers of the label.

Most experts (6/7) believed that it would be hard to fit all the purposes of data collection on the primary layer of the label. Therefore, they recommended including this information on the secondary layer. However, past research has shown that purpose of data collection is one of the most important factors consumers want to consider when making privacy decisions [50, 202, 207]. Purposes may be grouped into high-level categories that could be included on the primary layer. For example, the W3C's Platform for Privacy Preferences (P3P) standard identified 12

purpose categories [343]. The consumers we interviewed indicated that it was important to them to know the type of data collected and the purpose of collection when making device purchase decisions.

Experts stated that where the data is stored should be included on the secondary layer because it is not relevant to privacy or security. However, we believe local storage versus data being stored on the cloud can indeed have different privacy and security implications [296]. Therefore, we decided to include this in the primary layer. Moreover, most consumer participants were able to reason about privacy and security implications of cloud versus device and discuss the trade-off between security and convenience. C10 talked about this trade-off by saying:

The advantage of the cloud is that if the device is damaged, you can still access it. So it's going to be always available as long as you can access internet. The other issue with the cloud though is that, like, it can be hacked and also, who has access to that is less clear, or you have less control over that. But I can always access it from whatever device I have and it's convenient.

Although experts recommended including information about device resource usage on the secondary layer, we decided to remove this factor from the label due to its lack of privacy, security, or safety implications.

Factors with no Specific Layer

There were four remaining factors that at least 4 out of 7 experts were enthusiastic about having on the label, but their opinions were split between including them on either the primary or the secondary layer. These factors were:

- Who the data is shared with
- Who the data is sold to
- Whether or not the device can still function when data-driven smart features (e.g., the learning function of smart thermostat) are turned off
- Whether or not the device has parental control mode

For these factors, about half of the experts reported that they would like to include them on the primary layer since they are important privacy and security factors that consumers should know about before making purchase decisions. The other half of the experts were not enthusiastic about including these factors on the primary layer.

We decided to put the factors related to parental controls and device functionality when smart features are turned off on the secondary layer because they are not directly related to security or privacy.

Some experts (3/7) noted that who data is shared with or sold to is likely to change over time, and recommended putting these factors on the secondary layer where they could be updated more easily. We showed consumers a label with these factors on the secondary layer. However, all consumer participants expressed concern when they saw that their information could be shared and sold with third parties and 8 out of 15 said who data is sold to or shared with were among the most important factors that could inform their purchase decisions. Therefore we moved these factors back to the primary layer.

Factors to Exclude from the Label

There were six factors that at least 6 out of 7 experts believed should not be included on the label:

- List of device-compatible products
- Link to the device’s software and hardware bill of material
- Link to the device’s accompanying app(s)
- Whether or not the device manufacturer has a bug bounty program
- Where and when the device brand was incorporated
- Consumer Reports rating

The most common reasons experts said these factors were not suitable for the label were the lack of relevance to privacy and security and inability to convey risk to consumers. For example, S2 did not want the label to include whether or not the manufacturer has a bug bounty program as this factor does not offer adequate insight into security practices of the company: “This information is not too important on how the company does security analysis.”

Almost all experts were opposed to including the Consumer Reports rating, mainly as they believed this organization’s reputation does not stem from their privacy and security assessments. However, Consumer Reports is in the process of developing a digital privacy and security standard [322], so this may change in the future.

We decided to include whether or not the device manufacturer has a bug bounty program on the secondary layer. Since the word “bug bounty” was not immediately clear to consumer participants, we changed the wording to vulnerability disclosure and management, which was more understandable to them. When we presented this factor to our participants, 13 out of 15 associated this information with having good privacy and security practices, hence they were more inclined to trust the company who were transparent about their devices’ discovered vulnerabilities and had taken steps to manage them. C3 explained: “This factor shows that this is a company that has a security process, and participates in public activities to educate the community on things that can go wrong.” C5 wanted the IoT companies to disclose their devices’ history of known vulnerabilities: “A lot of times, if the company had some kind of vulnerability, they maybe want to sweep it under the rug and not let anyone know about it. That’s good that it shows you that they’re being honest about what issues they’ve had in the past, and what they’ve done to address them.”

We decided to include a link to the software and hardware bill of materials (mentioned on the label as software and hardware composition list) on the secondary layer since it can provide useful information related to security when it is available. When we asked consumers about this factor, most wanted it to be included and noted that even if they did not understand it, it could be useful to those with technical expertise.

Based on our experts’ opinions, we initially excluded the list of device-compatible products from the label. In the consumer interview study, we asked participants to tell us about anything they thought was missing from the label that they would like us to add. The only factor that participants suggested was a link to the privacy statement of device-compatible products such as Alexa. As this was suggested by a couple of our early participants, we added a factor on the secondary layer to list compatible platforms with a link to their privacy policies.

6.2.3 Attitudes toward Labels and Layered Design

All consumer participants discussed how difficult it is for them currently to find information related to privacy and security of smart devices before purchasing them. They all reported that they would like to have an IoT security and privacy label available at the point of sale, mainly to be as informed as possible.

Most experts mentioned that IoT privacy and security labels are useful to inform consumers when making purchase decisions, which is in line with our previous work, described in Chapter 5. P7 explained that a label can provide consumers with information they would not have otherwise:

What's good about a label is that it empowers the consumer to make a more active decision about cybersecurity rather than just being completely helpless as to what the security of her device might be. Especially as more and more of this technology is designed for consumers, the average consumer doesn't have a privacy, security, or a legal department to review this stuff before they buy it. Enterprises do, but consumers do not, so someone's gotta be looking out for consumers and giving the consumers this information.

All consumer participants were familiar with layered labels, as they had seen QR codes on products such as food, drugs, or video games. Most of our participants (11/15) expressed positive attitudes toward the layered design, mainly because the amount of information that could fit in two layers would not fit on a single-layer label. Participants also appreciated the ability to easily gain further insights about the privacy and security practices of the device manufacturer. These participants reported that they engage in a combination of online and in-store shopping. Hence, they believed the layered label design would be useful to them throughout their purchase process.

Some consumer participants (4/15) thought a layered label would not be ideal, citing the inconvenience of using a phone to scan the QR code when shopping in a store, especially for the elderly. C15 explained: "These technologies for older generation, they are kind of tough. The idea of installing something to scan the QR code, it's going to be too much for them. I know they prefer to just read everything, put on their glasses and read everything line by line." Two participants expressed concern that companies might withhold important information from the primary layer and put it only in the secondary layer.

For each of the factors on the label, we asked consumer participants to tell us how they believe that factor would impact their purchase decisions and whether they would like us to remove the factor or add additional details to the factor. All participants understood the factors presented on the primary layer and were able to discuss the potential risks associated with all of these factors except the level of detail for data storage. Although they all understood the terms "identifiable" and "anonymous," participants did not associate identifiable data with risk in this context, perhaps because the utility of a security camera is increased if it can record videos in which people are identifiable. Further testing is needed to understand the impact of the interaction between the purpose of data collection and the granularity of data on consumers' privacy concerns.

As we expected, some of the information on the secondary layer did not convey privacy and security risks to consumer participants. However, participants still asked to see most of the factors that we included on the secondary layer because they wanted to be as informed as possible when purchasing a smart device.

Participants mentioned that they might search online for information about unfamiliar factors and the availability of our label would help them. C5 explained: “I don’t know what TCP and UDP are. But it’s interesting to have this here, because then I could go to Reddit and ask on there what that means and what the capabilities are.”

There were only six factors that 2 or 3 consumer participants thought should be removed from the label. These secondary-layer factors were perceived by those participants as lacking relevance to privacy and security (physical actuators, hardware and software bill of material), not understandable (MUD compliant, key management protocols, open network ports), and not relevant to them (special data handling practices for children).

At least one consumer participant recognized that some of the factors in the secondary label might be useful to experts. C3 explained:

Labels are both for customers and for experts such as tech journalists, consumer advocacy groups, who are capable of understanding it and who will click on the things, and if they see something that is questionable will raise it in the public press, will raise it with regulatory authorities, and otherwise. The label is not just for the consumer, but also there’s another feedback process that works through experts to the extent that the information is available at all.

6.2.4 Prototype Privacy and Security Label

We used the results from our expert elicitation and consumer studies as well as recent IoT security standardization and certification efforts to inform the design of a prototype privacy and security label for a hypothetical smart security camera. Our design has primary and secondary layers, as shown in Figures 6.3 and 6.4, respectively. The secondary layer includes plus signs next to each item that can be clicked to reveal further details. We envision that the secondary layer would be accompanied by a computer-readable version of the label to enable automated processing and comparison between products, for example by personal privacy assistants [74] or search engines [80, 113]. Our website at www.iotsecurityprivacy.org has the latest label design.

6.3 Discussion

Expert participants recommended including privacy and security star ratings on an IoT label, mostly because ratings would help consumers to more easily compare IoT devices based on their privacy and security practices. All of our consumer participants liked the idea of having privacy and security assessments from trustworthy organizations. Although we believe these third-party assessments would inform consumers’ purchase behavior, we decided not to include them on our proposed label, as there is no organization currently doing these evaluations at scale for a wide range of IoT devices. We expect to add a place for assessment information once it is available.

We begin this section by providing a comparison between certifications and star ratings. Next, we discuss possible approaches to IoT privacy and security certifications.

Security & Privacy Overview

Smart Security Camera, NS200

Firmware version 2.5.1: updated on: 6/15/2019

The device was manufactured in: United States

Casa

 Security Mechanisms	Security updates	Automatic (available until 1/1/2022)	
	Access control	Password, Factory default, User-changeable, Multiple user accounts are allowed	
 Data Practices	Sensor data collection	 Video	 Audio
	Purpose	Providing device functions, research	Providing device functions, research
	Data stored on device	Identified	Identified
	Data stored on cloud	Identified, Option to delete	Identified, Option to delete
	Shared with	Manufacturer	Manufacturer
	Sold to	Not sold	Not sold
	Other collected data	Presence, Temperature, Carbon monoxide, Usage information, User-entered information	
	Privacy policy	www.NS200.example.com/privacypolicy	
 More Information	Detailed Security & Privacy Label: www.iotsecurityprivacy.org/labels		

Figure 6.3: Primary layer of the label (as of September 2019). This layer is designed to be printed on product packaging or to appear on a product website. View our latest label design at www.iotsecurityprivacy.org.

Security & Privacy Details

Smart Security Camera, NS200
 Firmware version 2.5.1, updated on: 6/15/2019
 The device was manufactured in: United States



<p>Security Mechanisms</p>	Security updates	Automatic (available until 1/1/2022)		+		
	Access control	Password, Factory default, User-changeable, Multiple user accounts are allowed		+		
	Security oversight	Audits performed by internal security auditors		+		
	Ports and protocols	www.NS200.example.com/port				
	Hardware safety	www.NS200.example.com/hwsafety				
	Software safety	www.NS200.example.com/swsafety				
	Personal safety	www.NS200.example.com/usersafety				
	Vulnerability disclosure and management	www.NS200.example.com/vulreport				
	Software and hardware composition list	www.NS200.example.com/BOM				
	Encryption and key management	www.NS200.example.com/key				
<p>Data Practices</p>	Sensor data collection	<p>Video</p>	<p>Audio</p>	<p>Presence</p>	<p>Temperature</p>	<p>Carbon Monoxide</p>
	Collection frequency	When user requests it	Continuous, Adjustable	Periodic, Option to opt-out	Continuous, Option to opt-in	Continuous, Option to opt-out
	Purpose	Providing device functions, research	Providing device functions, research	Providing device functions	Providing device functions	Providing device functions
	Data stored on device	Identified	Identified	De-identified	De-identified	De-identified
	Local data retention time	Up to a month	Up to a month	Up to a year	Up to a year	Up to a year
	Data stored on cloud	Identified, Option to delete	Identified, Option to delete	No cloud storage	De-identified	De-identified
	Cloud data retention time	Up to a month	Up to a month	No cloud storage	Up to a month	Up to a month
	Shared with	Manufacturer	Manufacturer	Not shared	Manufacturer, Third-party	Third-party, option to opt-out
	Sharing frequency	Periodic, Adjustable	Periodic, Adjustable	Not shared	Continuous	Continuous
	Sold to	Not sold	Not sold	Not sold	Third-party	Third-party, Option to opt-out
	Other collected data	Usage information, User-entered information				
	Data linkage	Data may be linked with internal and external data sources				
	What could be inferred from user's data	No data inference				
Special data handling practices for children	Yes					
In compliance with	GDPR, ISO27001					
Privacy policy	www.NS200.example.com/privacypolicy					
<p>More Information</p>	Call Casa with your questions at	412-313-2793 (24/7 support)		+		
	Functionality with no internet	Limited functionality on offline mode		+		
	Functionality with no data processing	Limited functionality on dumb mode		+		
	Physical actuations and triggers	Device blinks when motion is detected		+		
	Compatible platforms	Amazon Alexa		+		

Figure 6.4: Secondary layer of the label (as of September 2019). This layer of the label can be accessed by scanning the QR code or typing the URL on the primary layer. View our latest label design at www.iotsecurityprivacy.org.

6.3.1 Star Ratings vs. Certification Levels

Similar to the Energy Star rating system managed by the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Energy (DOE) [268], the idea of star ratings has been proposed for IoT devices to help consumers make informed purchase decisions [67, 261]. In a hearing of the U.S. Senate Committee on Commerce, Science, and Transportation’s Subcommittee on Security [338], Senator Ed Markey suggested a 5-star security rating system for IoT products. In our study, while experts were supportive of privacy and security ratings on the label, they also mentioned two potential challenges of including them.

The first challenge experts brought up relates to the rating scale. Experts suggested that consumers might have trouble distinguishing a large number of ratings, yet a more granular scale could help manufacturers better differentiate their privacy and security practices. P1, who works in industry, discussed this issue: “I’m sure industry people, manufacturers, will want more in there. What would happen if you had something like this is it might start to grow based on features they want reflected in that rating. Then I can see it becoming a bigger and bigger scale.”

Experts mentioned that ratings might pose an unhealthy incentive for IoT companies to achieve full-star ratings only to be able to compete in the market. Companies may be able to game the ratings in order to get all the stars and eventually all products will have all stars, whether they deserve them or not. S2, an academic, explained: “The problem I have with ratings like this is that everybody’s gonna get a five star, because everybody’s gonna figure out how to get the five star.”

To address these challenges, some experts discussed the idea of having multiple certification levels (e.g., silver, gold, platinum) with a secure baseline or minimum standard instead of star ratings. This is similar to what the LEED standards use for rating energy efficiency and sustainability of buildings [199]. P8 explained: “I think consumers should know it passes the minimum security level. If I’m buying a space heater, I know they’re not allowed to sell me one that will set on fire. I don’t have to say, oh, it has a 70% score that it will set the house on fire.”

Underwriters Laboratories (UL) published a 5-level IoT security standard (bronze, silver, gold, platinum, and diamond) in 2019 [333].

As of January 2020, no devices have been certified [334]. As manufacturers start having their devices certified, this certification could be added to the IoT label.

Since the lowest certification level indicates a safe device, there is a risk that manufacturers will aim to achieve the lowest level and not bother pursuing higher levels. Market competition may encourage manufacturers to pursue higher certification levels, especially for devices where the consequences of security breaches are most severe.

6.3.2 Privacy and Security Evaluation and Scoring

Over the past few years, a number of organizations and research teams have started to develop standards for IoT privacy and security evaluation and scoring. They include Consumer Reports [322], YourThings [14, 358], and UL [333].

Digital Standard

In 2017, Consumer Reports launched the Digital Standard to work toward providing a comprehensive standard that enables organizations to evaluate consumer IoT products. This standard focuses on four categories: security, privacy, ownership, and governance & compliance [322].

The security category of the Digital Standard includes build quality, data security, and personal safety.

Build quality refers to product stability and whether “software was built and developed according to the industry’s best practices for security.” The Cyber Independent Testing Lab (CITL) [85], a Digital Standard partner, is actively evaluating and scoring software of IoT devices according to a number of factors. Our label design includes a software safety features element where manufacturers can provide a URL with information related to software security.

Data security includes authentication, encryption, updatability, security audits, and vulnerability disclosure program. All of these factors are included on our label.

The personal safety category has not yet been defined in the Digital Standard, although developer notes indicate it will be related to avoiding abuse and harassment. Media reports suggest there are many incidents involving smart home devices being used for domestic abuse [52]. However, device manufacturers appear to be doing little to address the risks associated with abuse involving their devices [319]. We have included a factor called personal safety, which provides a place where device manufacturers can indicate available safeguards against abusive behavior once such safeguards have been implemented. Further discussions with experts are needed to determine how to address significant safety issues effectively on the label. As it was explained by S4: “Safety means if your car gets hacked, you die. The room that has a laser attached and if it gets hacked, it kills you. A drone can be reprogrammed to dive-bomb your child. I’m not sure how to capture that on the label.”

The privacy section in the Digital Standard includes user controls, data use and sharing, data retention, and overreach. The assessment procedure for almost all the privacy factors in the Standard involves verifying the company’s claimed data practices with actual data practices.

All the privacy factors mentioned in the Digital Standard are covered in our proposed label, except overreach. Overreach, or “collecting too much data” focuses on determining whether data collection is beneficial to the user, fully disclosed, the minimum necessary for functionality, and private by default. This seems like an area where a third-party assessment rather than a self report is likely warranted.

As some of the experts we interviewed mentioned, consumers may weigh privacy and functionality trade-offs differently. Thus it may be difficult to capture a single privacy rating that makes sense for all consumers. In addition to providing detailed information about data practices, a future privacy rating system could be customized based on a consumer’s stated privacy preferences, which could change over time.

YourThings

Alrawi et al. [14] developed a security evaluation and scoring method for smart home devices. In their YourThings [358] initiative, they produce device scorecards with grades in four areas: device, mobile application, cloud endpoints, and network communication. While our label pro-

vides information related to all of the major areas of the YourThings rubric as well as some security and privacy factors not addressed by YourThings, the YourThings scorecard considers some additional security details, including some that do not lend themselves to self report. The YourThings scorecard offers a concise expert summary of device security issues, which could be useful to include on the label.

The YourThings rubric considers five device-related factors: upgradability, exposed services, vulnerabilities, configuration, and Internet pairing. We include all on our label.

The mobile section of the YourThings rubric includes sensitive data, programming issues, and “over-privileged,” i.e., requesting excess permissions that are not used or required. Sensitive data is defined in the rubric to include “artifacts like API keys, passwords, and cryptographic keys that are hard-coded into the application.” Our label includes factors related to sensitive data and programming issues such as software safety features and key management protocol. However, over-privileged is a factor better assessed by a third-party evaluator rather than being self-reported.

The cloud endpoints section of the rubric includes domain categories, TLS configuration, and vulnerable services. Some of the information needed to compute this score is included on our secondary layer of the label when fully expanded. However, details needed to evaluate TLS configuration are not included.

Finally, the network communication section of the rubric includes protocols, susceptibility to Man in the Middle (MITM) attack, and use of encryption. While the secondary layer of our label provides some of the information needed to compute this score, a third-party evaluation would be needed to provide a complete assessment.

The concise YourThings scores are useful for comparing devices, but users may need to drill down to obtain information relevant to their specific needs. For example, devices are penalized for not having automatic updates. While automatic updates are generally considered the most secure approach, poorly timed updates can be problematic, potentially interfering with critical device functions.

UL

The 5-level UL certification process includes 44 requirements over seven categories: software updates, data & cryptography, logical security, system management, customer identifiable data, protocol security, and process & documentation [332].

While our proposed label includes factors from all seven categories, a third-party evaluation is needed to assess compliance with requirements. Our label can inform consumers about security and privacy, and goes into more detail about privacy issues than UL’s customer identifiable data category. By including the UL certification in our label, we would offer users a single concise assessment of device security that complements the more detailed information provided on the label.

6.4 Conclusion

We conducted a study with 22 privacy and security experts to elicit their opinions on the contents of IoT privacy and security labels. By following a three-round Delphi method, we found the factors that experts believed should be included on the label, and distributed them between primary and secondary layers of the label in three categories (security mechanisms, data practices, and general/more information). By conducting a series of in-depth semi-structured interviews with 15 IoT consumers, we iteratively improved the design of our proposed privacy and security label for IoT devices. The latest version of the label and implementation information, as well as a tool for generating the label are available at www.iotsecurityprivacy.org.

To design an effective privacy and security label, we need to make sure that the presented information effectively conveys risks to a larger set of representative consumers and that they are able to use it in their purchase decisions. In the next chapter, we present the results of a large-scale online study in which we assess the impact of each of 16 attributes on risk perception and willingness to purchase.

Chapter 7

Which Privacy and Security Attributes Most Impact Consumers' Risk Perception and Willingness to Purchase IoT Devices?

In previous two chapters, we conducted both expert and user interviews and surveys to provide a detailed proposal for what information should be included on an IoT security and privacy label. After interviewing a diverse set of privacy and security experts, we specified 47 privacy and security attributes that should be included on a two-layer IoT nutrition label. In addition, we conducted a small-scale interview study with 15 IoT consumers to observe whether participants understand the information presented on the label. Our final proposed design was, however, based primarily on the opinion of experts, who have been shown to often perceive risks differently than the regular public [356].

To complement the previous study and bridge the gap between experts' knowledge and consumers' understanding, we next conducted a large-scale survey study on MTurk with 1,371 participants. We presented participants with three randomly selected vignettes describing an IoT device purchase scenario with a label on the device describing a single privacy or security attribute. We tested the most protective and least protective values of 16 attributes specified in Chapter 6. After each scenario, we asked questions to capture respondents' perception of risk and their willingness to purchase, as well as the reasons behind their preferences.

We ranked the significance of each privacy and security attribute-value pair in describing participants' risk perception and willingness to purchase. We found that among all attribute-value pairs, those which explicitly or implicitly conveyed to users that their information could be shared with other parties significantly elevated perceived risks. On the other hand, those factors which conveyed that either no information is being retained or no information is being shared significantly increased respondents' willingness to purchase the device.

Our analysis indicates that participants are in most cases well-informed about the potential privacy and security risks and their consequences and although the perceived risk significantly

This chapter is a lightly edited version of a paper currently under submission as: Pardis Emami-Naeini, Janarth Dheenadhayalan, Yuvraj Agarwal, Lorrie Faith Cranor. Which Privacy and Security Attributes Most Impact Consumers' Risk Perception and Willingness to Purchase IoT Devices?.

influences their willingness to purchase, these two attitudes are not always perfectly aligned. Participants reported that some attributes reduced their perception of risk, but would not increase their desire to purchase the device.

We make the following contributions in this chapter:

- Through our quantitative data collection, we identify the privacy and security attributes and corresponding values that most impact participants’ risk perception and willingness to purchase IoT devices.
- Through our qualitative data collection, we gain insights into why participants were influenced by label attributes to perceive a device as more or less risky or to report being more or less likely to purchase that device.
- We distill a set of actionable recommendations on how to better inform consumers’ purchase behavior by more effectively conveying privacy and security risks to consumers in an IoT label format.

7.1 Methodology

We conducted a large-scale vignette study [127] on MTurk with 1,710 participants (reduced to 1,371 participants after filtering the responses) from the United States who were at least 18 years old. On average, it took participants 20 minutes to complete our survey, and we paid them \$2.50 for their time.

7.1.1 Study Design

To explore the impact of privacy and security information on participants’ IoT-related risk perception and willingness to purchase, we designed our study with two between-subject factors—the *device type* and the *recipient of the device*. We tested two types of devices and three types of device recipients for a total of six experimental conditions. Our within-subject factor was the IoT-related privacy and security information conveyed on the label. Each participant was randomly shown 3 of the 33 possible pairs of attributes and their corresponding values.

We randomly assigned each participant to one experimental condition. They received survey questions all related to the device type and recipient associated with their condition. The survey questions are provided in Appendix E.

Prior to launching the main study, we piloted our survey on MTurk to look for potential misunderstandings and determine how long it would take participants to answer our questions. We found that by presenting participants with three attributes, the survey completion time would be on average 20 minutes. Therefore, to mitigate survey fatigue [27], each participant was asked to answer the survey questions for three randomly-chosen attribute-value pairs.

After presenting participants with the consent form and CAPTCHA verification, we started the survey by asking about the participants’ concern level and purchase history for a specific type of smart device. We then presented participants with three vignettes about the purchase of an IoT device, using the device type and recipient in their assigned condition. Each vignette included mention of a product label with a single attribute-value pair, selected randomly. To enhance par-

ticipants' information comprehension, we provided an explanation next to each attribute-value pair. Appendix E.1 shows the explanations we used in the survey.

We asked participants how the information on the label would change their risk perception and their willingness to purchase and the reasons behind their assessments. We also asked an attention-check question related to each attribute. After asking about the three vignettes, we presented participants with the complete list of privacy and security attributes, specified in Chapter 6, and asked them to specify whether they were interested in knowing more about each of the attributes and whether having this additional information would change their willingness to purchase the device. We then asked a question to capture participants' understanding of how their assigned smart device collects data. We concluded the survey by asking a number of demographic questions. Participants were then shown a random code, which they had to type into the MTurk portal to complete the survey.

Between-Subject Factors

Concerns and attitudes change risk perception [132, 304, 305] and IoT device privacy concerns are impacted by the types of data collected [202, 203, 256, 362]. Therefore, we considered device type as a between-subject factor and tested two types of devices: smart speaker (with a microphone that will listen and respond to voice commands) and smart light bulb (with a presence sensor that detects whether someone is present in the room to control the lighting automatically). We chose these devices to represent two extremes of perceived privacy concerns. A smart light bulb has been shown to raise few concerns [117], while a smart speaker has been shown to raise many privacy concerns, due in part to the fact that it captures voice data [270, 362].

Our other between-subject factor was the IoT device recipient. The risk target has been shown to be an effective factor in influencing people's risk perception [303, 304]. We were interested in understanding whether participants have different risk perceptions and willingness to purchase based on who they are purchasing the device for. Therefore, we tested three conditions: purchasing the device for yourself, gifting it to a family member, or gifting it to a friend.

Concern Level and Purchase History

We asked participants to specify how concerned they were about the smart device collecting data and the reason for their answer. Next we asked them if they currently have a smart device of that type in their home. If participants had the smart device, we asked them how long ago they purchased it, whether they own the device and how they acquired it, what brand they purchased, and why they purchased the device. If they did not have the smart device, we asked them whether they were ever in the market to purchase it and if so, we asked them what made them decide not to purchase it.

Privacy and Security Label Attributes

We selected 16 of the 47 privacy and security attributes identified in Chapter 6. Because many of the 47 attributes simply specified URLs that linked to information provided by the device manufacturer (e.g., the value for software safety might be `www.NS200.example.com/swsafety`),

Layer	Attribute	Tested value	
		Most protective	Least protective
Primary	Security update	Automatic	None
	Access control	Multi-factor authentication	None
	Purpose	Device function	Monetization
	Device storage	None	Identified
	Cloud storage	None	Identified
	Shared with	None	Third parties
	Sold to	None	Third parties
	Control over	Cloud data deletion Device storage	
Secondary	Average time to patch	1 month	6 months
	Security audit	Internal & external	None
	Collection frequency	On user demand	Continuous
	Sharing frequency	On user demand	Continuous
	Device retention	None	Indefinite
	Cloud retention	None	Indefinite
	Data linkage	None	Internal & external
	Inference	None	Additional info
	Control over	Device retention	

Table 7.1: The 16 security and privacy attributes along with the values of each attribute tested. The attributes are partitioned to the ones included in the primary and secondary layers of our previous proposed IoT label (see 6.3 and 6.4). The additional info value of the attribute inference was described as “characteristics and psychological traits, attitudes and preferences, aptitudes and abilities, and behaviors” in the survey.

we tested only the 16 attributes that had a set of discreet values.

Based on our review of IoT privacy and security standards and guidelines, we synthesized the possible values each attribute might take and identified the most protective and least protective values of each attribute to test in the study. For one of the attributes we considered three values (the complete list of the attribute-value pairs are presented in Table 7.1). Out of these 33 attribute-value pairs, we randomly selected three to present to each participant in the form of a vignette describing a hypothetical purchase scenario. Each vignette took the following form:

Imagine you are making a decision to purchase a [*device type*] for [*device recipient*]. This device has a [*device sensor*] that will [*device data collection*]. The price of the device is within your budget and the features are all what you would expect from a [*device type*]. On the package of the device there is a label that explains the privacy and security practices of the [*device type*].

The label on the device indicates the following: [*attribute: value*]

Our overarching research goal was to specify whether the label attributes impacted participants' risk perception and willingness to purchase in the expected directions, i.e., most protective decreases risk and increases desire to purchase (Table 7.1). The secondary goal was to recommend improvements to the IoT label, e.g., by identifying common misconceptions that require further explanation.

Level of confidence has been identified as an influential factor to explain risk perception [78, 156, 308]. To be able to calibrate participants' responses based on confidence level, we asked how confident they were that they knew what the information presented meant.

To understand participants' risk perception, we asked respondents to specify how the presented privacy and security attribute-value impacts the privacy and security risks associated with the device in question. Participants could choose from "Strongly decreases," "Slightly decreases," "No impact," "Slightly increases," "Strongly increases," or "Not sure." We then asked participants to explain the reason behind their choice.

We asked similar questions to understand the impact of the privacy and security attributes on participants' willingness to purchase the device for the recipient based on the condition they were assigned to.

To check our participants' attention to the survey questions, we tested participants on the privacy and security information they were answering questions about with a multiple choice question. For example if the presented *attribute-value* was *Security audit: none*, we asked participants "Which statement is correct about the device described in the previous question?" and provided three incorrect answers and the correct answer "The manufacturer does not conduct security audits on its devices and services."

Additional Privacy and Security Attributes

To get a more holistic sense of participants' level of interest and willingness to purchase, we presented them with additional questions in a matrix format. In these questions, the rows corresponded to the complete list of the privacy, security, and general attributes from Chapter 6 (47 attributes total, split over three matrices) and the columns were "interested to know about" and "would impact my willingness to purchase the device." Participants could check or uncheck each box to indicate their interest or willingness.

Perceived Device Functionality

To understand how participants perceived the device data collection, we asked them to choose whether they believe the device is always sensing data, collecting data only when it is triggered (e.g., by mentioning the wake word or by someone turning on the light), collecting data only when a user pushes a physical button on the device, or does not expose its data collection methods.

Demographic Questions

At the end of the survey, we asked general demographic questions to capture participants' age, gender, highest degree earned, and background in technology if any.

7.1.2 Data Analysis

We conducted a mixed between-subjects and within-subjects study to understand the impact of various factors on participants' risk perception and willingness to purchase. We applied Generalized Linear Mixed Model Regression (GLMM) to find the best models that describe our dependant variables (DV) of willingness to purchase and risk perception. Since we conducted a repeated-measure study, we used random intercept in the models to count for within-participants data dependencies. GLMM is particularly useful when modeling a repeated-measure design, in which participants are presented with multiple parallel scenarios with same type of questions [40].

To construct the models that best describe our DVs, we used Bayesian Information Criterion (BIC) as the fit metric and applied backward elimination. We started by including all the factors we were interested in, including within-subject and between-subject factors, device ownership, ordering of the vignettes, and five demographic factors. In each step of the model selection, we removed the factor that had the highest p -value and calculated the BIC again, until BIC reached its global minimum, which indicates the current set of the factors best describes the DVs [176]. We picked the significance threshold of 0.05 to specify statistically significant findings. We present the regression results of the model selection in the Results section 7.2.

The survey responses that capture participants' risk perception and willingness to purchase were on a 5-point Likert scale. However, to construct our mixed-effect logistic regressions, we grouped the responses into a binary factor of 0 and 1. In models, where the DV was the perceived risk or the willingness to purchase, the responses were coded as 0 if they were "slightly decreases," "strongly decreases," or "having no impact." The DVs were coded as 1 if the responses were "slightly increases" or "strongly increases."

To analyze free-text responses, the first author was the primary coder, who constructed the codebook and kept it updated throughout the process. We conducted content analysis to find the most common reasons as to why participants' risk perception and willingness to purchase were impacted or not impacted by privacy and security information [291].

Two researchers independently applied the codebook to the free responses and iteratively revised the codebook. After resolving the coding disagreements, we reached the Cohen's Kappa inter-coder agreement of 81%, which is considered an *excellent* rate of agreement [131]. In case of disagreement, we report on the findings of the primary coder.

7.1.3 Limitations

In this study, we tested the impact of privacy and security attributes on participants' *self-reported* risk perception and willingness to purchase. These self-reported assessments provide useful insights about how participants understand and evaluate these attributes. Although risk perception and willingness to purchase have been shown to strongly correlate with actual risk and purchase decision [9, 269, 280, 342, 360], to assess actual influence on purchasing behavior future work should observe real purchase scenarios.

In our study we could only evaluate the importance of a limited number of factors in describing risk perception and willingness to purchase. For instance, we only tested two types of IoT devices, three types of recipients, and two values of most security and privacy attributes.

Future studies should consider additional factors that could potentially influence risk perceptions and willingness to purchase such as cultural differences [88, 89, 107], price [264], and social proof [58].

We randomly presented each participant with three privacy and security attribute-value pairs and asked questions about each. Participants' risk perception and willingness to purchase related to the second or third attribute-value pair could be biased due to their answers to the preceding questions. To test for the ordering effect, which has been shown to impact risk judgment [57], we included an ordering factor in the risk-perception and willingness-to-purchase models. By changing the baseline of this factor, we found evidence of bias in responses to the third willingness-to-purchase question. We did not find evidence of bias in responses to the risk-perception questions. Therefore, we only report results for the first two willingness-to-purchase questions, but report results for all three risk-perception questions.

7.2 Results

We initially recruited 1,683 MTurk participants and excluded those whose answers for all our open-ended questions were irrelevant. This resulted in 1,371 participants who are included in our analysis. All of these participants answered at least two out of their three attention-check questions correctly. Overall, at least 90% of participants correctly answered the attention-check questions for all but two of the 33 attribute-value pairs, indicating that participants were paying attention to and had at least a basic understanding of the label information we presented to them. The two attribute-value pairs with most wrong answers were: *control over: device storage* (21% incorrect), and *security audit: internal & external* (22% incorrect).

Our participants were 54% male and 45% female. Compared to 2018 US Census [336] data, participants were younger and better educated. Participant demographic information is provided in Table 7.2.

In this section, we start by discussing results of our quantitative analysis of concern level and purchase history, followed by risk perception and willingness to purchase. Next, we provide insights from our qualitative analysis into the reasons behind participants' responses. We then talk about the most common metrics participants used to assess the risks. Finally, we provide a discussion on assessing the risk perception and willingness to purchase by testing the extreme values of each privacy and security attribute.

7.2.1 Concern Level and Purchase History

We queried our participants about their concerns related to data collected and used by the smart device. For the smart speaker conditions, 93% of the participants reported being concerned while only 62% of those in the smart light bulb conditions were concerned. Furthermore, our regression results showed that being concerned significantly increased participants' risk perception (p -value < 0.001). Concern has been previously identified as one of the factors influencing risk perception [132]. In contrast, level of concern did not have a significant impact on participants' willingness to purchase the device.

Metric	Levels	MTurk (%)	Census (%)
Gender	Male	54.0	48
	Female	45.5	51
	Non binary	0.005	—
Age	18-29 years	23.3	21.0
	30-49 years	61.3	33.3
	50-64 years	13.0	25.1
	65+ years	2.3	20.4
Education	No high school	0.2	10.9
	High school	28.8	47.1
	College	51.2	20.6
	Professional	10.6	11.6
	Associate	8.6	9.6
	No answer	0.3	—
Tech Background	Yes	19.8	—
	No	80.1	—

Table 7.2: Demographic information of our participants and 2018 US Census data [336]. In some cases, the Census data did not include a specific category, which we denote by —.

For the smart speaker condition, 54% of participants reported having a smart speaker in their home, and among those 53% purchased the device themselves. The most common reasons for purchasing a smart speaker were convenience and a desire to try new technology. Smart light bulbs were not as popular among our participants, with only 12% reporting having one, and 61% of those reporting purchasing it for themselves for convenience and security purposes. We applied Kendall’s tau correlation test and found that not having the smart device is strongly correlated with being concerned about that device (p -value < 0.05).

Among those who did not have the smart device in question, 23% reported that they were in the market to purchase it earlier. The main reasons stated for not going through with the purchase for both devices were their price (30% speaker and 48% light bulb) and the lack of necessity (44% and 34%). Privacy and security concerns were reported by 26% and 9% of participants as reasons not to purchase the smart speaker and the smart light bulb respectively.

7.2.2 Models to Describe Risk Perception and Willingness to Purchase

We were interested in understanding the impact of various factors on two dependent variables (DV): participants’ risk perception and willingness to purchase. The summary statistics in Tables 7.3 and 7.4 show how participants specified their risk perception and willingness to purchase the smart device.

The information in Tables 7.3 and 7.4 does not consider the within-participants data dependencies. To count for such dependencies, we built two mixed models to describe our DVs. The factors we initially included in each model are as follows.

Attribute-value	% Increased risk	% Decreased risk	% No impact	% Not sure
Cloud retention: none	0.81	81.14	11.47	6.55
Access control: MFA	1.57	79.52	11.02	7.87
Cloud storage: none	1.65	70.24	7.43	20.66
Control over: device retention	2.40	67.20	10.40	20
Device retention: none	3.12	80.46	5.46	10.93
Shared with: none	3.14	81.88	3.93	11.02
Data linkage: none	3.22	62.09	14.51	20.16
Control over: device storage	3.90	78.90	3.90	13.28
Inference: none	3.90	42.18	24.21	29.68
Data collection: on user demand	6.06	65.90	7.57	20.45
Sold to: none	7.31	69.10	3.25	20.32
Sharing frequency: on user demand	8.66	62.99	8.66	19.68
Control over: cloud data deletion	9.01	66.39	4.91	19.67
Security update: automatic	11.71	57.03	7.03	24.21
Device storage: none	14.63	46.34	10.56	28.45
Security audit: internal & external	31.20	37.60	13.60	17.60
Purpose: device function	43.90	12.19	13.82	30.08
Security audit: none	55.37	3.30	34.71	6.61
Inference: additional info	67.22	0.84	21.00	10.92
Average time to patch: 1 month	73.77	13.11	5.73	7.37
Security update: none	74.40	1.60	14.40	9.60
Cloud storage: identified	79.52	3.14	6.29	11.02
Purpose: monetization	81.45	0.80	6.45	11.29
Data linkage: internal & external	81.81	0.82	11.57	5.78
Collection frequency: continuous	82.92	1.62	3.25	12.19
Device storage: identified	84.00	3.20	8.00	4.80
Sharing frequency: continuous	86.06	1.63	7.37	4.91
Average time to patch: 6 months	87.02	3.05	4.58	5.34
Cloud retention: indefinite	87.90	0.00	2.41	9.67
Device retention: indefinite	89.07	0.00	3.36	7.56
Shared with: third parties	92.00	0.80	3.20	4.00
Sold to: third parties	92.12	0.00	5.51	2.36
Access control: none	93.65	0.00	3.96	2.38

Table 7.3: Summary statistics showing percentages of risk perception for each attribute-value sorted by the “Increased risk” column.

- `sp_attribute_value`: 33 attribute-value pairs (see Table 7.1).
- `device_type`: Type of the device, with two levels: Smart speaker and smart light bulb.
- `device_recipient`: Who the device is being purchased for, with three levels: Purchasing for yourself, gifting to a friend, and gifting to a family member.
- `device_ownership`: How the participants acquired the smart device, with three levels: Having purchased the device themselves, owned the device but did not purchase it, and did not have the device.
- `order`: The order of the vignette presented to participants, with three levels: First, second, and third.

Attribute-value	% Increased willingness	% Decreased willingness	% No impact
Shared with: none	73.33	5.33	21.33
Device retention: none	66.66	3.84	29.48
Sold to: none	63.85	9.63	26.50
Cloud retention: none	63.09	4.76	32.14
Control over: device storage	56.52	8.69	34.78
Collection frequency: on user demand	53.84	13.18	32.96
Control over: cloud data deletion	53.76	6.45	39.78
Access control: MFA	53.48	6.97	39.53
Control over: device retention	53.08	3.70	43.20
Data linkage: none	51.94	2.59	45.45
Sharing frequency: on user demand	47.19	11.23	41.57
Device storage: none	46.57	15.06	38.35
Cloud storage: none	44.31	3.40	52.27
Inference: none	40.00	4.00	56.00
Security update: automatic	38.66	9.33	52.00
Security audit: internal & external	30.00	26.25	43.75
Purpose: device function	18.18	31.81	50.00
Average time to patch: 1 month	10.84	65.06	24.09
Average time to patch: 6 months	7.14	74.48	18.36
Security update: none	6.89	73.56	19.54
Device storage: identified	6.75	74.32	18.91
Sharing frequency: continuous	5.19	72.72	22.07
Cloud storage: identified	3.57	77.38	19.04
Collection frequency: continuous	3.48	76.74	19.76
Data linkage: internal & external	2.29	74.71	22.98
Cloud retention: indefinite	2.10	77.89	20.00
Inference: additional info	1.33	73.33	25.33
Security audit: none	1.25	66.25	32.50
Device retention: indefinite	1.21	69.51	29.26
Sold to: third parties	1.17	89.41	9.41
Shared with: third parties	1.17	87.05	11.76
Purpose: monetization	1.14	75.86	22.98
Access control: none	1.08	88.04	10.86

Table 7.4: Summary statistics showing percentages of willingness to purchase for each attribute-value sorted by the “Increased willingness” column.

- Demographic information: Age, gender, education level, and whether they have a background in technology (see Table 7.2).

In the regression analysis, the significance of the levels within each attribute should be compared to the baseline of that attribute. We selected *purpose: device function* to be the baseline for `sp_attribute_value`, as it is the purpose that most IoT devices will have, possibly in addition to others. The *smart light bulb* is selected as the baseline for `device_type`, since its data collection is less concerning, and the baselines for factors `device_ownership` and `device_recipient` are selected to be the most common values of these factors.

Level of concern with the device also had a significant impact on the risk perception (p -value < 0.001). However as we previously mentioned, this factor also had a strong correlation with `device_ownership`. Therefore, we could not include both of these factors in the regression

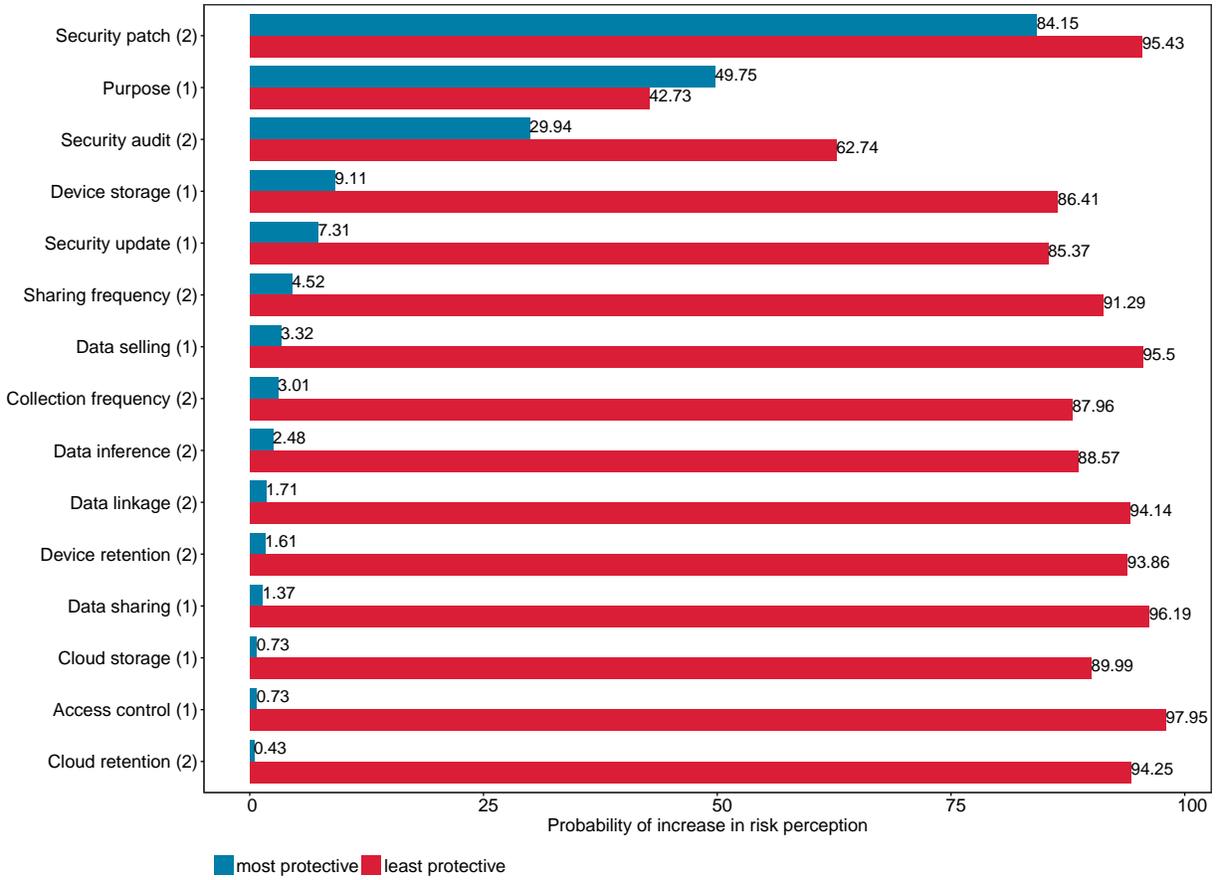


Figure 7.1: The probability of increase in the perceived risk for the most protective and least protective values of attributes. The primary layer attributes are denoted by (1) and the secondary layer attributes are denoted by (2).

model as the independent variables (IV). We included `device_ownership` in the model as it helped the model fit better (by looking at the Bayesian Information Criterion (BIC) of the model) compared to including the concern level.

We used GLMM with random intercept to build our models and applied backward elimination to find the most influential factors that best describe participants’ risk perception and willingness to purchase. The final regression models with the minimum global BICs are presented in Tables 7.5 and 7.6.

For the risk-perception model, a positive estimate for an attribute-value pair indicates that providing that information increases risk perception compared to the baseline. Similarly, for the model for participants’ willingness to purchase, a positive estimate for an attribute-value pair indicates that providing that information increases their willingness to purchase, and a negative estimate indicates hesitance to purchase.

In both models, all the privacy and security attribute-value pairs that we tested were aligned with our hypothesized risk level (see Table 7.1), except for the average time to patch. The Underwriters Lab (UL) guidelines suggest that the most severe vulnerabilities should be patched

Attribute-value	Estimate	Risk probability	Std Err	Z-value	p-value
(Intercept)	-0.01	49.75	0.27	-0.06	0.96
sp_attribute_value (baseline = purpose: device function)					
Cloud retention: none	-5.42	0.43	1.06	-5.10	0.00***
Access control: MFA	-4.90	0.73	0.80	-6.11	0.00***
Cloud storage: none	-4.90	0.73	0.81	-6.02	0.00***
Control over: device retention	-4.48	1.10	0.70	-6.41	0.00***
Shared with: none	-4.26	1.37	0.64	-6.66	0.00***
Device retention: none	-4.10	1.61	0.61	-6.68	0.00***
Data linkage: none	-4.04	1.71	0.63	-6.38	0.00***
Control over: device storage	-3.90	1.94	0.58	-6.76	0.00***
Inference: none	-3.66	2.48	0.60	-6.11	0.00***
Data collection: on user demand	-3.46	3.01	0.52	-6.72	0.00***
Sold to: none	-3.36	3.32	0.51	-6.63	0.00***
Sharing frequency: on user demand	-3.02	4.52	0.48	-6.33	0.00***
Control over: cloud data deletion	-2.97	4.83	0.47	-6.34	0.00***
Security update: automatic	-2.53	7.31	0.43	-5.83	0.00***
Device storage: none	-2.29	9.11	0.42	-5.40	0.00***
Security audit: internal & external	-0.84	29.94	0.37	-2.24	0.02*
Average time to patch: 1 month	1.68	84.15	0.40	4.20	0.00***
Security audit: none	2.27	90.55	0.48	4.68	0.00***
Cloud storage: identified	2.29	90.72	0.44	5.26	0.00***
Collection frequency: continuous	2.32	90.97	0.44	5.32	0.00***
Inference: additional info	2.33	91.05	0.46	5.08	0.00***
Purpose: monetization	2.52	92.48	0.45	5.57	0.00***
Security update: none	2.55	92.68	0.47	5.43	0.00***
Cloud retention: indefinite	2.89	94.68	0.48	6.07	0.00***
Average time to patch: 6 months	3.05	95.43	0.48	6.34	0.00***
Device retention: indefinite	3.06	95.47	0.50	6.11	0.00***
Device storage: identified	3.07	95.52	0.49	6.21	0.00***
Sharing frequency: continuous	3.14	95.81	0.51	6.12	0.00***
Data linkage: internal & external	3.15	95.85	0.52	6.04	0.00***
Shared with: third parties	3.70	97.56	0.57	6.53	0.00***
Access control: none	4.33	98.68	0.67	6.30	0.00***
Sold to: third parties	4.44	98.82	0.71	6.22	0.00***
device_type (baseline = smart light bulb)					
Smart speaker	0.55	63.18	0.17	3.26	0.00**
device_ownership (baseline = never have had the device)					
Not purchased the device	-0.42	39.41	0.23	-1.87	0.06
Purchased the device	-0.80	30.78	0.21	-3.74	0.00***
Observations	3735				
Log-Likelihood	-1054.6				
Akaike Inf. Crit.	2183.2				
Bayesian Inf. Crit.	2413.5				
Note: * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$					

Table 7.5: GLMM model to describe risk perception. A positive estimate indicates that the factor increases the perceived risk.

Attribute-value	Estimate	Purchase probability	Std Err	Z-value	p-value
(Intercept)	-1.57	17.22	0.33	-4.70	0.00***
sp_attribute_value (baseline = purpose: device function)					
Shared with: none	3.00	80.69	0.45	6.60	0.00***
Device retention: none	2.58	73.30	0.43	5.97	0.00***
Sold to: none	2.55	72.72	0.43	5.93	0.00***
Cloud retention: none	2.44	70.47	0.42	5.80	0.00***
Control over: device storage	2.03	61.30	0.43	4.74	0.00***
Collection frequency: on user demand	2.01	60.82	0.41	4.95	0.00***
Control over: cloud data deletion	2.01	60.82	0.40	4.96	0.00***
Access control: MFA	1.97	59.86	0.41	4.80	0.00***
Control over: device retention	1.93	58.90	0.41	4.67	0.00***
Data linkage: none	1.81	55.97	0.42	4.34	0.00***
Sharing frequency: on user demand	1.66	52.24	0.40	4.11	0.00***
Device storage: none	1.64	51.74	0.42	3.92	0.00***
Cloud storage: none	1.52	48.75	0.40	3.77	0.00***
Security update: automatic	1.24	41.82	0.42	2.97	0.00**
Inference: none	1.24	41.82	0.42	2.98	0.00**
Security audit: internal & external	0.83	32.30	0.42	1.98	0.05*
Average time to patch: 1 month	-0.70	9.36	0.49	-1.44	0.15
Average time to patch: 6 months	-1.16	6.12	0.52	-2.24	0.03*
Security update: none	-1.19	5.95	0.54	-2.19	0.03*
Device storage: identified	-1.26	5.57	0.58	-2.17	0.03*
Sharing frequency: continuous	-1.48	4.52	0.62	-2.41	0.02*
Collection frequency: continuous	-1.82	3.26	0.68	-2.69	0.00**
Cloud storage: identified	-1.83	3.22	0.68	-2.69	0.00**
Data linkage: internal & external	-2.29	2.06	0.79	-2.89	0.00**
Cloud retention: indefinite	-2.44	1.78	0.79	-3.09	0.00**
Security audit: none	-2.94	1.08	1.06	-2.77	0.00**
Device retention: indefinite	-2.94	1.08	1.06	-2.78	0.00**
Inference: additional info	-2.96	1.06	1.06	-2.78	0.00**
Shared with: third parties	-3.00	1.02	1.06	-2.83	0.00**
Purpose: monetization	-3.04	0.98	1.06	-2.87	0.00**
Sold to: third parties	-3.06	0.96	1.06	-2.89	0.00**
Access control: none	-3.14	0.89	1.06	-2.96	0.00**
device_type (baseline = smart light bulb)					
Smart speaker	-0.41	12.13	0.14	-2.87	0.00**
device_ownership (baseline = never have had the device)					
Not purchased the device	0.68	29.11	0.20	3.47	0.00***
Purchased the device	0.50	25.54	0.18	2.75	0.00**
device_recipient (baseline = for yourself)					
For a family member	0.11	18.84	0.15	-0.70	0.49
For a friend	-0.52	11.00	0.16	-3.28	0.00**
Observations	2742				
Log-Likelihood	-1079.9				
Akaike Inf. Crit.	2237.8				
Bayesian Inf. Crit.	2468.6				
Note: * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$					

Table 7.6: GLMM model to describe willingness to purchase. A positive estimate indicates that the factor increases participants' willingness to purchase the smart device.

within 1 month, less severe vulnerabilities within 3 months, and the least severe vulnerabilities could be possibly left unpatched [331]. Thus, we hypothesized that participants' perceived risk

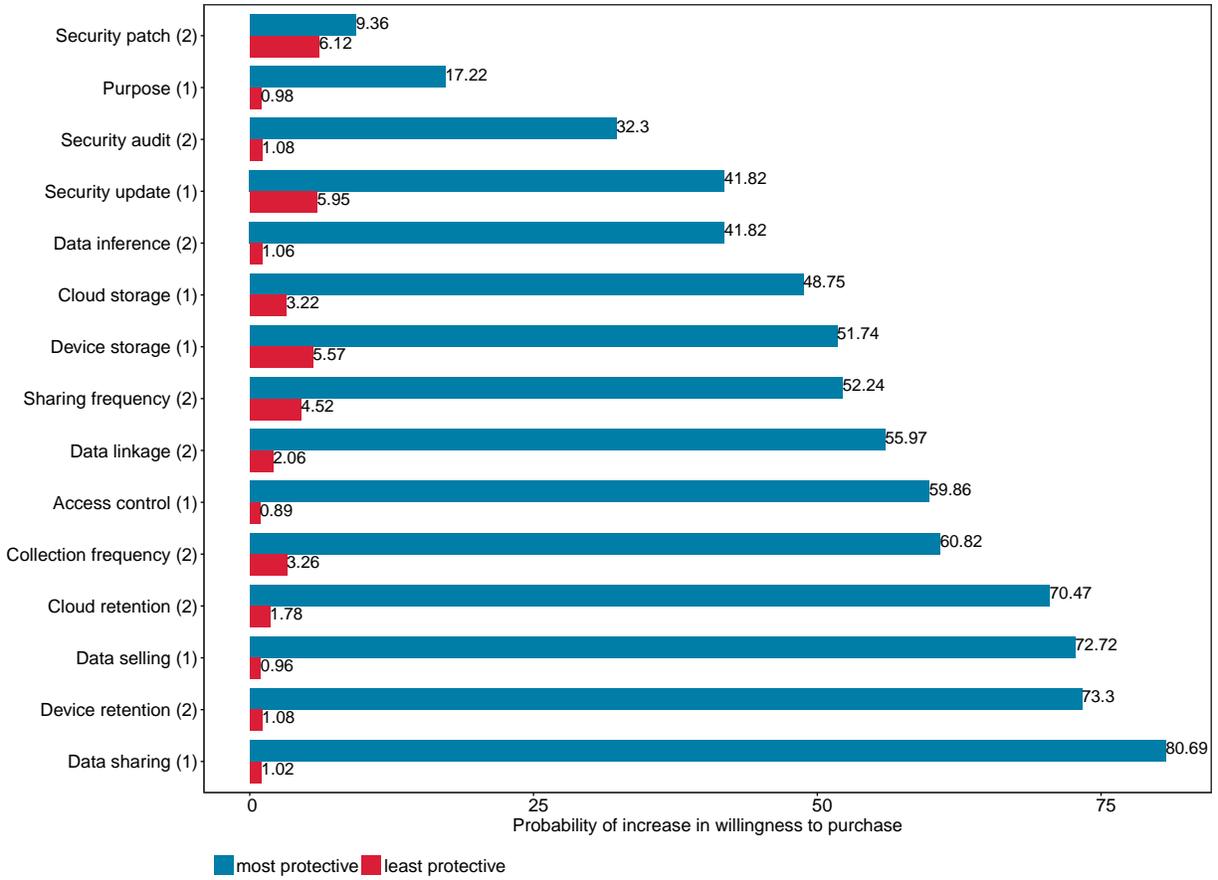


Figure 7.2: The probability of increase in willingness to purchase for the most protective and least protective values of attributes. The primary layer attributes are denoted by (1) and the secondary layer attributes are denoted by (2).

would decrease knowing that the vulnerabilities would be patched within 1 month, while a time to patch within 6 months would increase it. However, our regression results showed that average time to patch of both 1 month (estimate = 1.68, p -value < 0.001) and 6 months (estimate = 3.05, p -value < 0.001) strongly increase the perceived risk. In contrast, the 6-month patch period (estimate = -1.15, p -value < 0.05) strongly decreases participants’ willingness to purchase the smart device. The 1-month patch (p -value > 0.05) was not a statistically significant factor to describe willingness to purchase.

Of the 16 attributes in our regression models, 15 had two values, each expected to have an opposite impact. In Tables 7.5 and 7.6, based on the model estimates, we calculated the conditional probabilities of increase in risk perception and increase in willingness to purchase given each attribute-value pair. These probabilities are shown in Figures 7.1 and 7.2, respectively. From Figure 7.1, we can see that except for the extreme values of purpose of data collection, the least protective values of all other attributes caused participants to perceive a higher risk compared to their most protective values. We observe a similar trend in Figure 7.2, where the most protective values of all attributes increased participants’ desire to purchase the smart device.

As the figures show, average time to patch had the weakest impact on changing participants' risk perception and willingness to purchase. On the other hand, access control and cloud retention had the highest impact on risk perception, while device retention and who the data is shared with and sold to had the most impact on willingness to purchase.

Figure 7.3 shows a jitter (scatter) plot of participants' perceived risk levels and willingness to purchase when presented with the privacy and security attributes and their most protective and least protective values. As can be observed, the correlation between risk perception and willingness to purchase differ based on the attribute. For instance, Figure 7.3a shows that most participants perceived multi-factor authentication (MFA) as decreasing risk and no access control as increasing risk. While this generally corresponded with their willingness to purchase, the figure shows that some participants who perceived MFA as risk reducing were actually less likely to purchase a device with MFA. Likewise, Figure 7.3b shows that most participants perceived no sharing as decreasing risk and sharing with third parties as increasing risk. However, in this case risk perception was much more likely to be correlated with willingness to purchase. On the other hand, Figure 7.3p shows that participants perceived both values of average time to patch as risky and would decrease their likelihood to purchase. Finally, as we hypothesized, Figure 7.3h confirms that all levels of user control seem to generally reduce the participants' perceived risk and increase their willingness to purchase.

To figure out whether two values of an attribute have significantly different impacts on risk perception or willingness to purchase, we used the Wilcoxon Ranked Sum Test. This method is suitable here as we had a repeated-measure design. The assumption of this test is that the tested participant groups should be independent. Therefore, to test the values of each attribute, we first removed the participants, who had seen both values of an attribute and then conducted the analysis. Since we conducted the test for 15 of the attributes, we corrected for the p -value. The results showed that except average time to patch and purpose of data collection, the two values of each attributes had a significantly different impact (p -value < 0.003) on the risk perception and willingness to purchase.

In addition to the privacy and security attributes, the regression results showed that the `device_type` was a significant factor to describe both dependent variables. In particular, compared to a smart light bulb, participants perceived a strongly higher risk for a smart speaker (estimate = 0.54, p -value < 0.01) and they were significantly less willing to purchase the smart speaker (estimate = -0.41, p -value < 0.01). The `device_recipient` was not a statistically significant factor to describe risk perception. However, participants' willingness to purchase the device significantly decreased when the recipient of the device was a friend (estimate = -0.52, p -value < 0.01) compared to purchasing the device for themselves. The qualitative responses indicated the most common reason being the belief that friends should decide for themselves whether they are comfortable with stated privacy/security practices. The open-ended responses showed that participants felt more responsible when purchasing devices for themselves or family members compared to friends.

The last significant factor in the final models of risk perception and willingness to purchase was `device_ownership`. Previous work has shown that familiarity with the technology impacts risk perception [122, 135, 143]. Therefore, we tested whether participants would perceive a different level of risk if they currently owned the smart device and if they had also purchased the device themselves.

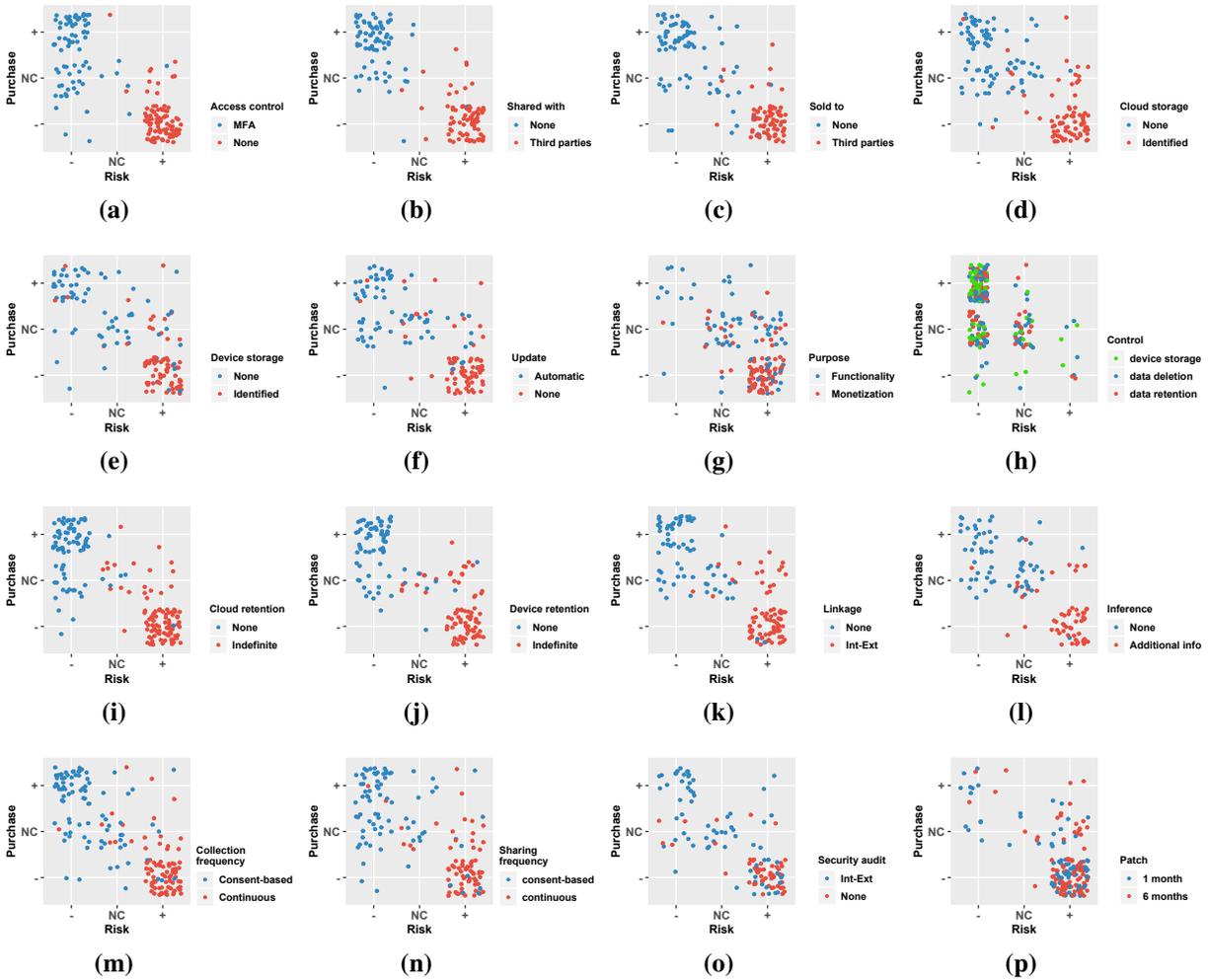


Figure 7.3: Jitter (scatter) plot of participants’ willingness to purchase vs. perceived risk for all 16 privacy and security factors and levels. For each metric, we have binned “strongly increase” and “slightly increase” (denoted by +), and “strongly decrease” and “slightly decrease” (denoted by –). No change in the risk perception and willingness to purchase is denoted by NC.

We found that those who had purchased the device themselves (estimate = -0.79, p -value < 0.001), perceived significantly less risk than those who have never had the device. We also found that compared to those who have never had the device, having it would significantly increase the willingness to purchase a future device. This includes those who had purchased the device themselves (estimate = 0.49, p -value < 0.01) and those who had the device, but received it in other ways (estimate = 0.67, p -value < 0.001). Nevertheless, we found no significant differences based on how participants acquired the device.

We also asked participants how confident they were about knowing what the presented privacy and security information meant. Among all attribute-value pairs, participants reported having the highest confidence about the meaning of *device retention: indefinite*, followed by *sold to: third parties* and *shared to: third parties*. They reported the least confidence in *audit: none*. We found that the level of confidence had a significant impact on participants’ risk perception.

Participants who had more confidence perceived potential risks significantly lower than those who had less confidence. This finding is aligned with risk literature on the impact of confidence and certainty [78, 156, 308]. The extent of confidence did not have a significant impact on the willingness to purchase.

7.2.3 Qualitative Results

Although participants' risk perception and willingness to purchase were generally aligned with our hypotheses, there were some participants who responded differently. By examining the open-ended explanations participants provided, we identified the most common reasons that some participants' risk perception and willingness to purchase differed from our hypotheses.

Insufficient Information

Wanting more information was the most common reason mentioned by participants who thought that an attribute-value pair would not have an impact on their risk perception and/or willingness to purchase. These participants reported that they would like to have more information to make an informed decision. These attribute-value pairs include:

- *Security audit: internal & external*: Participants wanted to know who the auditors were and what information they had access to when conducting the audits.
- *Device storage: none*: participants reported that they would like to know whether information would be collected on the cloud or if no device storage means no data will be collected in general.
- *Security update: automatic*: participants requested to know how often their device would get updated.
- *Average time to patch: 1 month*: A number of participants reported being unsure about whether one month was too short or long, expressing that they need more information on why it takes manufacturers one month to fix vulnerabilities.

Lack of Trust in Manufactures

The second most common reason as to why a factor would not have any impact on participants' risk perception and willingness to purchase was not trusting manufacturers to follow their stated practices.

Participants expressed lack of trust mostly when the purpose of data collection was for device functionality. Although we hypothesized that providing data for device functionality should decrease the perceived risk, that was true for only 12% of participants seeing that information. The other participants stated that this information would not impact their risk perception or would increase the risks due to their lack of trust in manufacturers. As P778 explained: "The companies who collect data are incredibly untrustworthy. They do not have consumers' best interests in mind when they are utilizing the data they collect."

A few participants mentioned lack of trust when assessing the risk perception of automatic updates. They reported that manufacturers can apply any unwanted changes to their devices

under the premise of security updates.

Shared with: none and *sold to: none* were other attribute-value pairs where participants expressed lack of trust. Some participants mentioned that while they believe this information would decrease the potential privacy risks with the smart device, they do not trust the manufacturers not to send their data to other companies for profit.

Participants' comments about trust are consistent with prior work that has identified the trust people have in organizations as one of the factors effecting risk perception [13, 84, 112, 242, 257, 300, 301, 330].

The Standard Process

For a number of attribute-value pairs, participants believed that the reported privacy and security practices were standard, and therefore having them would not provide additional privacy and security protection.

These pairs were *security update: automatic*, *collection frequency: on user demand*, *sharing frequency: on user demand*, *access control: multi-factor authentication*, *control over: device storage*, *control over: device retention*, and *control over: cloud data deletion*.

P878 believed that data collected with user consent is standard: "I would assume this is standard and normal. If the company is not ethical they will just collect the data anyway."

Usability Challenges

For a few privacy and security attribute-value pairs, participants mentioned that having such information would lead to difficulty in using the device. These pairs were *access control: multi-factor authentication*, *collection frequency: on user demand*, and *inference: none*.

Participants reported that requiring users to use additional authentication methods or consenting to data collection each time would affect device usability.

P1334 was particularly concerned about MFA for shared in-home devices: "Accessing the device via authentication would then become a hassle and/or annoying. For instance, what if my wife or a guest wanted to use the speaker?"

All participants who mentioned the usability challenges reported that these pairs would decrease the risks. However, they would also decrease their willingness to purchase.

Desire to have Control

Automatic security update is recommended by a number of IoT privacy and security guidelines [118, 172, 322, 331] and this *attribute value* significantly decreased participants' risk perception. However, for some participants this decrease did not change their willingness to purchase due to the lack of control implied by this practice.

P535 reported: "I want to have full control over updating my devices to decrease the risk of installing an update that has a security flaw."

No Initial Concerns

In the experimental conditions where the device was a smart light bulb, many participants exhibited low initial concern levels, mostly due to not seeing the consequences of the data collection. Therefore, privacy and security *attribute values* that generally reduced risk perceptions had no impact on some participants because they did not perceive a risk to begin with.

For instance P28 reported: “To me, the type of data that a smart bulb would collect does not seem to be consequential in relation to one’s personal privacy. If it strictly collects information based on motion detection, this isn’t a big concern.”

Misunderstandings

There were a few attribute-value pairs that some participants misunderstood, thus affecting their responses. One such pair was *security update: none*. Some participants mentioned that receiving no security updates implied maximum security protection, alleviating the need for updates. Another misunderstood attribute was the average time to patch. Some participants mentioned that a device that receives security patches must not be secure or it would not need patches. For instance, P906 mentioned: “On the label it advertises that patches are even needed. That is why there is a perception of decreased privacy.”

Other Reasons

There were some reasons that were mentioned by only a few participants. One of the reasons that a attribute-value pair did not change participants’ willingness to purchase was that they had already made their decisions to either purchase or not purchase the device, due to factors such as its functionality or their prior privacy and security concerns with the device, the latter of which was only mentioned by participants who were asked questions about the smart speaker.

P750, who was asked to imagine purchasing a smart speaker, reported: “There is little incentive for the companies to keep data secure. The fact that IoT devices all send data to a privately controlled server is unacceptable. Any low level employee or barely motivated hacker could get access to all the information. The government could just ask for the information. I don’t want any such devices in my home.”

Another reason that was mentioned by participants who were in the experimental conditions in which they were asked to imagine purchasing the smart device for a family member or a friend as a gift, was that they did not feel comfortable making a decision for the gift recipient without knowing their preferences.

The final reason that was mentioned by some participants was that the privacy and security attribute-value pair is not enough to eliminate potential risks of the device. For instance, P1338 reported: “Just because I can control how the data is retained on the device doesn’t mean I have any control about how that data is collected and how it is used while it is retained - the company could still upload the data from the device to their server before it is deleted from the device.”

Attribute-value	Data	Time	Visibility	Protection
Cloud retention: none	—	↓	—	—
Access control: MFA	—	—	↓	↑
Cloud storage: none	↓	—	—	—
Control over: device retention	—	↓	—	—
Shared with: none	—	—	↓	—
Device retention: none	—	↓	—	—
Data linkage: none	↓	—	—	—
Control over: device storage	↓	—	—	—
Inference: none	↓	—	—	—
Collection frequency: on user demand	↓	—	—	—
Sold to: none	—	—	↓	—
Sharing frequency: on user demand	—	—	↓	—
Control over: cloud data deletion	↓	—	—	—
Security update: automatic	—	—	—	↑
Device storage: none	↓	—	—	—
Security audit: internal & external	—	—	—	↑
Average time to patch: 1 month	—	—	—	↓
Security audit: none	—	—	—	↓
Cloud storage: identified	↑	—	—	—
Collection frequency: continuous	↑	—	—	—
Inference: additional info	↑	—	—	—
Purpose: monetization	↑	—	↑	—
Purpose: device function	↑	—	—	—
Security update: none	—	—	—	↓
Cloud retention: indefinite	↑	—	—	—
Average time to patch: 6 months	—	—	—	↓
Device retention: indefinite	—	↑	—	—
Device storage: identified	—	↑	—	—
Sharing frequency: continuous	—	—	↑	—
Data linkage: internal & external	↑	—	—	—
Shared with: third parties	—	—	↑	—
Access control: none	—	—	↑	↓
Sold to: third parties	—	—	↑	—

Table 7.7: Breakdown of participants’ risk perception criteria using four metrics of data, time, visibility, and protection. ↑ indicates that the *attribute value* would increase the metric, ↓ indicates that the *attribute value* would decrease the metric, and — indicates that the *attribute value* would not have an impact on the metric. The green symbols indicate reduced risk perception, while the red symbols indicate increased risk perception.

7.2.4 Decision Criteria to Assess Risk

Based on participants' answers to open-ended questions on why a *attribute value* changes their risk perception, we found four primary decision criteria participants used to assess the increase and decrease in risk due to a privacy and security attribute-value pair. These criteria are the amount of personal information (*data*), amount of time the information is available (*time*), the number of people who have access to the information (*visibility*), and the amount of protection (*protection*). Participants referred to protection as having another layer of security. As shown in Table 7.7, if an *attribute value* reduces any of the first three criteria or increases the last one, that attribute-value pair would decrease participants' perceived risk and vice versa. *Security audit: internal & external* was the only polarizing attribute-value pair: there was a disparity between participants who thought audits would increase the level of security protection of their device and participants who associated external auditors with third parties and were concerned about their data being shared with them (suggesting a misunderstanding about security audits).

7.2.5 Impact of the Extremes

We investigated the impact of two extreme values of each privacy and security attribute on participants' risk perception and willingness to purchase. We found that for all attributes except two (the average time to patch and the purpose of data collection), the change in both risk perception and willingness to purchase from one extreme value to another is statistically significant.

Nevertheless, these extremes were not the only levels that each attribute could exhibit. Indeed, there were a few cases, where participants provided us with insights into how they would perceive other values of attributes. For instance, when assessing the perceived risk of having vulnerabilities patched within 1 month or 6 months, some participants reported that they would like the vulnerabilities to be patched within one or two days of discovery.

An attribute for which participants mentioned values other than the extremes was security updates. When asked to assess the risks of having automatic updates, about half reported that security updates would decrease the risks of the device. However, they preferred to give their consent before updates were installed (consent-based updates [14, 358]).

7.3 Discussion

We begin this section by examining our label design proposed in Chapter 6, in light of our data on which attributes best communicate risk and influence decision making. We then propose recommendations to more effectively inform consumers' purchase behavior.

7.3.1 Label Content

Our proposed label has two layers (see Figures 6.3 and 6.4). The distribution of attributes between layers and the order of the attributes on each layer is based on experts' risk perceptions.

Although our findings on the importance of each attribute in this chapter have some overlap with our previously proposed label design, there are some differences that are worth highlighting.

Current Attributes on the Primary Layer

Among the seven attributes that we include in the primary “Overview” layer, intended for consumers to view on product packaging, our results indicated that access control, shared with, and sold to are also included in the top-seven most effective factors impacting risk perception and willingness to purchase. In Figures 7.3a-7.3c, we can see a clear separation between risk perception and willingness to purchase for the extreme values of these attributes.

Our analysis showed that cloud storage is among the top-seven most influential factors for risk perception. However, this attribute had little impact on participants’ reported desire to purchase the smart device (see Figure 7.3d), mainly due to not understanding the adverse consequences of cloud storage.

Device storage was not among the top-seven most influential attributes for either risk perception or willingness to purchase. About half of the participants viewed no device storage as a privacy-protective practice, but the rest inferred that no device storage implies storing the data on the cloud (a misconception). Since participants were generally concerned about identified data storage on the device, we believe this attribute should stay on the primary layer. However, information related to device storage should be communicated to consumers more clearly.

Security updates and the purpose of data collection had little influence on risk perception or willingness to purchase. Our participants were concerned with the least protective values of these attributes and perceived a significantly higher risk. However, the most protective values did not prove to be effective in changing their willingness to purchase the device.

We found that participants understood the importance of receiving security updates (see concentration of red dots in Figure 7.3f). However, they were not impressed by automatic updates, mainly due to a desire to control updates and a lack of trust in manufacturers issuing updates.

When the purpose of data collection was monetization, 81% of participants reported that it would lead to more risk. Yet, they were not convinced that providing device functions was much better, as only 12% of them thought it would decrease risk. The most commonly stated reason was that the description of *providing device functions* was too vague, and they associated that vagueness with manufacturers’ attempt to collect users’ data for other purposes, raising concerns. To improve risk communication, IoT manufacturers should provide more detailed information on the secondary layer of the label about what the device functions are and how the collected data supports device functionality. Security audit was another attribute that several participants reported that they would like to see more information about before making the device purchase. This additional information should be included on the secondary layer of the label.

Attributes to be Added to The Primary Layer

Cloud retention (Figure 7.3i), device retention (Figure 7.3j), and data linkage (Figure 7.3k) were not on the primary layer of our previously proposed label (see Figure 6.3). Nonetheless, we believe this information should be included on the primary layer as it impacts participants’ risk perception and willingness to purchase the device. As shown by our regression results (see

Table 7.5), *cloud retention: none* was the single most effective *attribute value* in decreasing risk perception. Participants were also significantly concerned about data being linked to internal and external data sources.

7.3.2 Information Presentation

Our quantitative and qualitative findings indicated that overall, presenting privacy and security attributes on an IoT label would influence participants' risk perception and impact their willingness to purchase the device. Although almost all of the tested attribute-value pairs were statistically significant in explaining risk perception and willingness to purchase, we found a number of ways that privacy advocates and IoT manufacturers can better communicate risk to consumers and help inform their purchase behavior.

More Information and Less Uncertainty

As previously mentioned, one of the commonly reported reasons as to why an attribute would not impact a participant's risk perception or willingness to purchase was not having enough information to make an informed decision. Participants also reported uncertainty about how the described privacy and security *attribute value* would harm them.

As we reported, having more confidence that they understood privacy and security information significantly decreased participants' risk perception, consistent with prior risk literature [78, 156, 308]. Providing consumers with more or clearer information would decrease their uncertainty, which could lead to perceiving less risk (unless, of course, the information makes it clear that the device is indeed risky).

To help consumers make more informed decisions, IoT manufacturers should provide them with specific details about each attribute without overwhelming them. To accomplish this, we propose adding extra information on the secondary layer of the label in an expanded view accessed by a plus sign that is placed next to each attribute.

Make Control Usable

Usability and desire to have control were two common reasons as to why providing the most protective attributes could potentially reduce consumers' willingness to purchase the device. Our findings showed that the perceived usability challenges of attributes such as multi-factor authentication would make participants more hesitant to purchase the smart device. It is important to note that for all the *attribute values* that raised usability-related concerns, participants also agreed that the *attribute value* would decrease the risks of the device.

We found that having control over three types of data practices would decrease the perceived risk and increase the willingness to purchase the device (see Figure 7.3h). Although the majority of our participants specified that automatic security updates would decrease risk, this information did not impact their willingness to purchase, mostly due to the lack of user control implied by the factor. Aligned with prior work [285], our participants preferred knowing about the details of each update before allowing installation. Although having control was favorable for some attributes, continuously asking for users' consent, on the other hand, would lead to usability

challenges. For instance, participants believed that asking the user to consent to data collection would decrease the risks, but it would also be a barrier to using the device.

Considering both usability and the desire to have autonomy, IoT manufacturers should provide users with choices about the level of control they would like to have over their devices and provide convenient interfaces for exercising that control. Moreover, since the ability to control has been shown to decrease the perceived risk [95, 241], IoT manufacturers need to clearly convey the potential risks and benefits of each of the offered choices to bridge the gap between the perceived risks and actual risks [294, 327]. Due to the significance of user controls in changing consumers' risk perception and willingness to purchase, the availability of such controls should be provided on the primary layer of the label and any additional information about these controls should be presented on the secondary layer of the label.

7.3.3 Viewing Label as a Whole

In this project, we explored the changes in risk perception and willingness to purchase caused by extreme values of various privacy and security attributes. We found that the most protective and least protective values of attributes significantly influence participants.

Our envisioned IoT label would include several privacy and security attributes on the same label. Hence, it is important to test the risk communication of the label when attributes are included in unison. To achieve this, a future factorial or fractional-factorial study could be conducted to test the significance of the combination of the label attributes.

7.4 Conclusion

Consumers are not aware of the privacy and security practices of their smart devices, and this lack of knowledge at the time of purchase could expose them to privacy and security risks. One possible solution to better inform consumers' purchase decisions is to provide privacy and security information on a label, similar to a nutrition label for food products. As described in Chapter 6, we conducted an expert elicitation study to specify the content of the label and proposed a layered privacy and security label. The focus there was to identify the most important privacy and security information to present on the label. We conducted only preliminary testing with consumers.

In this chapter, we conducted a mixed-design study with a larger set of 1,371 MTurk participants to measure the information efficacy along two dimensions: risk perception and willingness to purchase. The within-subject factor was the privacy and security information and the between-subject factors were the device type and the recipient of the device. In each experimental condition, each participant was randomly assigned to assess the risk perception and willingness to purchase related to three privacy and security attribute-value pairs. We also recorded participants' reasons behind their assessments.

Overall, we found that our label attributes successfully conveyed risk to participants. We found that participants' risk perception was more strongly influenced by label information than their willingness to purchase. This observed difference in influence was mostly due to lack of information, usability challenges, lack of trust in IoT manufacturers, desire to have control over

privacy and security, and lack of initial privacy and security concerns with the smart device. Based on our study findings, we proposed a number of recommendations to IoT manufacturers to more effectively convey risks to IoT consumers. These suggestions include providing more complete information about security and privacy, and providing consumers with choices about the desired level of control over their privacy and security.

Chapter 8

Conclusion and Future Work

In this chapter, we first provide a brief summary of the projects previously discussed in the thesis. We then briefly discuss the effectiveness and usefulness of having privacy and security labels for IoT devices. We will then review the labeling schemes proposed by the governments of UK, Finland, and Singapore. Next, we will introduce the specification document that fully describes our privacy and security label and the interactive online tool we have designed to generate our label. We conclude this section by discussing how the label can be further improved in future work and how we envision the label to be adopted.

8.1 Summary of the Discussed Research

In this thesis, we started by exploring how users of Internet of Things (IoT) devices make decisions related to data collected by them in different scenarios. We specified various factors that were significantly effective in explaining users' IoT-related preferences and expectations. Being surrounded by others, individuals often consider social cues when making decisions. We tested the impact of social cues from privacy experts and friends on privacy decisions related to IoT data collections. We found that people follow social cues from experts and their friends, especially when those cues present a large consensus and are consistent with their own opinions.

In the context of IoT, another common decision users make is whether or not to purchase a device. Consumers are increasingly purchasing IoT devices, but it is less clear whether they know much about the privacy and security of these devices at the time of purchase. To understand the importance of privacy and security in IoT consumers' purchase decision making, we conducted interviews and surveys with consumers of IoT devices and found that privacy and security are among the factors that consumers consider when purchasing a smart device. Yet as they are unable to find sufficient information about the privacy and security practices of the device, they make their purchase decision with limited relevant knowledge and information.

Consumers' interest in knowing more about the privacy and security attributes of IoT devices that they intend to purchase motivated us to design a tool to convey this information to consumers in a usable and effective fashion. To this end, we worked toward designing an informative label that covers the most critical privacy and security information consumers need to know about when purchasing IoT devices.

While numerous IoT standards and guidelines exist, they are almost all targeted toward IoT manufacturers or regulatory bodies. Furthermore, handful of them written to inform consumers and even fewer have done any user testing to know whether consumers can understand their documents. To fill this gap, we conducted a series of studies and systematically gathered input from a diverse group of privacy and security experts from industry, academia, government, and public policy organizations. In particular, we asked them to specify the most important privacy and security information consumers should know about.

Based on experts' input, we identified 47 attributes to include on our label. To increase the readability of the label, we prioritized the attributes and designed a layered label consisting of primary and secondary layers. The primary layer of the label is the concise format of the label including only a few critical privacy and security attributes. The secondary layer of the label is in an online-only format, has more attributes with additional information, and could be accessed from the primary layer by typing in a URL or scanning a QR code. We conducted a small-scale user study and iteratively enhanced the wording of the content of the label to make sure consumers have a basic understanding of the meaning of attributes and their values.

To evaluate the effectiveness of the content of the label in conveying risk to consumers and impacting their purchase behavior, we conducted a large-scale user study and specified the most effective and least effective privacy and security attributes and values. Among other findings, we found that participants perceive the highest risk when they know that third parties have access to their data or when they cannot specify who can access their device. Participants' willingness to purchase the device was highest when they were told that their data would not be shared with anyone.

We found that the usability of privacy and security practices is an important factor in people's willingness to purchase IoT devices. For example, although multi-factor authentication was perceived to significantly reduce the risks associated with the IoT device, due to its usability challenges, many participants reported that they are unwilling to purchase a device having this feature. Therefore, to further improve the label's risk communication, we provided a number of recommendations, including reducing the uncertainty of the information and the potential consequences by providing additional information about privacy and security practices and enhancing their usability.

8.2 On the Usefulness of Labels

Labels have been widely used to increase consumers' awareness in various domains, including energy and nutrition. However, in addition to the content of the label, the actual impact of the labels on consumers' purchase behavior depends on various factors, some of which are related to the consumers and some are related to the information presentation of the label. These include personal factors, such as level of knowledge and motivation in processing the label information [189, 251, 317], and socio-demographic factors, such as age, gender, family size, and income [72, 233, 250, 258, 346]. Label wording and formatting have also been shown to be effective factors in predicting the impact of nutrition label on consumers' purchase behavior. For example, using generalized claims as well as promotional claims on the label can lead to significant nutritional misunderstanding [18, 61].

Although consumers are increasingly worried about the privacy and security of their smart devices [159], some consumers might not have enough motivation to purchase smart devices with better privacy and security practices, regardless of the amount of information being provided to them at the point of sale. As we previously mentioned, literature on food nutrition label has shown that having prior nutritional knowledge significantly impacts people’s use of nutrition labels. From the prior work, we also know that knowledge predicts motivation [249] and motivation predicts knowledge [248]. Using this knowledge in the context of IoT, we hypothesize that providing knowledge to some consumers could initiate a positive cycle of knowledge and motivation, which could then lead to changing consumers’ purchase behaviors by means of IoT labels. Media reports have increased the much-needed awareness among consumers of IoT devices about the devices’ privacy and security practices, which could incentivize them to seek for more knowledge and information when purchasing such devices. To further motivate consumers to use the labels, future work should look into ways to improve consumers’ knowledge on the privacy and security implications of IoT devices.

8.3 International Labeling Efforts

Acknowledging the significance of labeling smart devices in informing consumers’ purchase behavior, other governments have started looking into labeling smart devices. Specifically, governments of the UK, Finland, and Singapore have recently published proposals on their labeling schemes.

8.3.1 United Kingdom

In 2018, the UK government published the Code of Practice for Consumer IoT Security [100], which includes 13 guidelines considered as good security practices for IoT manufacturers to follow. Later in 2019, the European Telecommunications Standards Institute (ETSI) published the Technical Specification 103 645, the first globally-applicable industry standard for consumer IoT security based on the Code of Practice guidelines [120].

In particular, three of the aforementioned guidelines were leveraged by the UK government to specify three attributes to include on the label. These three guidelines are for the manufacturers to have unique passwords for their products, to have a vulnerability disclosure policy, and to specify the end date for the device to receive security updates [99]. Due to their significance, on our label we included access control, vulnerability disclosure and management, and security updates to cover the UK-specified guidelines. We provided the information on access control and update lifetime on the primary layer and a link to the vulnerability disclosure and management policy on the secondary layer. The UK proposed label is shown in Figure 8.1.

8.3.2 Finland

Finland is the first European country to certify smart devices to increase consumers’ awareness of devices’ security practices at the time of purchase [128]. To receive the “Security Badge,” the Finnish Transport and Communications Agency Traficom requires IoT manufacturers to fill out



Figure 8.1: The IoT security label proposed by the UK government.

and submit the security compliance form. Currently in their pilot program, the products of three IoT companies have received the security badge: Cozify Hub for smart homes, DNA’s Wattinen smart heating system, and the Polar Ignite fitness smartwatch.

The requirements mentioned in the compliance form are based on the standard issued by ETSI [120]. The attributes in the compliance form are: the availability of timely and signed security updates, the lifetime of software updates, list of certifications and regulations the device has complied with, access control, having vulnerability disclosure program, the average time to patch the vulnerability, what personal data is being collected, how data is being collected, the purpose of data collection, who has access to the data, where the data is being stored, information on encryption and key management, and information on ports, protocols, and services and how they are secured. Our label covers all the attributes listed in the compliance form. Figure 8.2 shows the The Finland’s security badge awarded to Polar Ignite fitness smart watch.

8.3.3 Singapore

In Singapore, the Cyber Security Agency (CSA) is working on introducing the Cybersecurity Labeling Scheme (CLS) for IoT devices in 2020. To allow IoT manufacturers and the market to adjust, the labeling will be rolled out as a voluntary program. When launched, Singapore will be the first Asia-Pacific region to introduce IoT security labels. CSA’s main goals to label IoT products are to help consumers make informed purchase decisions and at the same time incentivize IoT manufacturers to develop and provide products with enhanced security practices [321].

In the beginning, CLS will focus on the routers and smart home hubs as they are often the



Figure 8.2: The IoT security badge awarded to Polar Ignite fitness smartwatch¹ by the Finnish Transport and Communications Agency Traficom.

gateways into the rest of the home network. However, CSA’s plan is to have a scheme broad enough so as to cover a broad range of consumer IoT devices in the future [32].

Although the details of CLS have not been announced as of March 2020, CLS is expected to comprise various levels of cybersecurity ratings to inform consumers when purchasing IoT devices. These ratings will be based on metrics of no default password and software safety features (e.g., not having common vulnerabilities in the software and being resistant to penetration testing) [321]. Our label provides information on all the metrics specified by the CSA.

8.3.4 Other International Activities

To further improve the security of IoT devices at the international level, in July 2019, the Homeland Security and Public Safety Ministers of Australia, Canada, New Zealand, the UK, and the US agreed to work toward enhancing the security by design for consumer IoT devices and engage other nations to do the same [101]. In addition, the partner countries (France, Uruguay, UK, Canada, Senegal, Japan, US, and New Zealand) in the IoT Security Platform suggested nine common principles to consider while developing international frameworks. Some of these principles are to ensure having security updates for the device with a specified minimum length of support, requiring unique credentials, encrypting the data in transit and at rest, enabling easy data deletion for users, protecting personal information, and implementing vulnerability disclosure policy [170].

¹<https://www.polar.com/fi/ignite>

8.4 Specification and Tool Accompanying the Label

To provide further clarification on the label and help consumers and manufacturers understand the information provided on the label, we prepared a specification and designed a tool to help manufacturers easily generate the label for their products.

8.4.1 Specification for Privacy and Security Label

In addition to designing the label, we prepared an extensive specification to accompany the label. For each attribute in the label, we specify the values and sub-attributes the attribute can take, other references which mention the attribute, additional information that manufacturers can potentially provide, and best practices related to the attribute.

To prepare the specification document, we looked into more than 70 IoT privacy and security references from industry, non-profit organizations, government agencies, and academia. We also looked into international efforts in labeling IoT devices from the UK [100], Finland [128], and Singapore [321] that we discussed in Section 8.3. All the label attributes have been mentioned at least once in other references. On average, each security attribute has been mentioned by 20 other references. However, the average number of references per attribute is much lower for privacy attributes with an average of only 5 references per attribute. This huge difference shows how little current standards and guidelines discuss privacy practices of IoT devices compared to their security mechanisms.

The security attributes with the highest number of references are security updates (44 references), encryption and key management (38 references), ports and protocols (28 references), and vulnerability disclosure and management (26 references). The only three security attributes, which were mentioned by fewer than 10 references, were software and hardware composition list (7 references), personal safety (3 references), and security oversight (1 reference). Although these attributes were not as highly referenced as the rest of the security attributes, we believe they provide valuable information for consumers (personal safety) and experts (software and hardware composition list, and security oversight).

As previously mentioned, privacy attributes were not as highly referenced as security attributes. There were nine privacy attributes that were mentioned by at least five references: Who the data is being shared with (14 references), purpose of data collection (10 references), the granularity of the data stored on the cloud (10 references), privacy policy (9 references), sensor data collection (8 references), local data retention time (7 references), sensor type (6 references), data stored on device (6 references), and cloud data retention time (5 references). The rest of the privacy attributes, including frequency of data collection and data sharing, were only mentioned by one or two references. Although who the data is being shared with was the highest mentioned privacy attribute (14 references), who the data is being sold to was only mentioned by Online Trust Alliance (OTA) [273]. From our study on measuring the effectiveness of the label information in risk perception and willingness to purchase (Chapter 7), we found that knowing whether the data is being sold to third parties has the highest impact on participants' perceived risk and reported desire to purchase the device. The latest version of the specification can be found at www.iotsecurityprivacy.org.

8.4.2 Tool to Generate the Label

To enable manufacturers to generate our labels and help standardizing the label design, we created a label generator wizard. This tool allows users to generate the labels by filling out a form and selecting the appropriate values for each privacy, security, and general attribute. As users are filling out the form, they can see the label being updated in real time. At any point, users can download the JSON format of the label and work on it locally. The JSON file can be uploaded at anytime to resume working on the label. In addition to JSON, the tool lets users download the XML and HTML formats of the label as well. Figure 8.5 shows first page of the tool, including the instruction to generate the label as well as the staged process to complete the sections of the label. The most current version of the tool can be found at www.iotsecurityprivacy.org².

Label Example

We looked into the information provided by the manufacturer Ring, who was itself acquired by Amazon, to generate a label for their Ring Doorbell, using our label generator. It is no surprise that the manufacturer has not disclosed a large number of the attributes on the label. There were some attributes that were only mentioned without any useful details provided (e.g., software safety, encryption and key management). Based on the data explicitly mentioned in the company's documents, we designed the labels in Figures 8.3 and 8.4.

8.5 Future Directions to Enhance the Label

In the rest of this section, we outline several potential directions to further improve the labels.

8.5.1 Design Elements of the Label

In this thesis, we discussed a number of quantitative and qualitative studies to design and evaluate a privacy and security label for IoT devices. In all the conducted projects, we focused mostly on improving the content comprehension of the label and only briefly looked into how information should be presented on the label.

From risk literature, we know that to maximize risk communication, it is of great importance for the risk communication method to trigger people's attention [55]. The literature on nutrition label has emphasized the importance of the label formatting on changing consumers' purchase behavior. An important future direction, therefore, is to focus on the design elements of the label, including but not limited to the amount of information, order of the attributes on each layer, the font size and color, and the order of the sections. All these design elements could impact consumers' risk perception and potentially their purchase behavior.

²Special thanks to Shreyas Nagare for the development of the website and the tool.

Security & Privacy Overview



Ring Video Doorbell 2

Firmware version: Not disclosed

The device was manufactured in: United States

 Security Mechanisms	<table border="0"> <tr> <td>Security updates</td> <td colspan="3">Automatic</td> </tr> <tr> <td>Access control</td> <td colspan="3">Password - user generated - user changeable, multi-factor authentication, single user account is required</td> </tr> </table>				Security updates	Automatic			Access control	Password - user generated - user changeable, multi-factor authentication, single user account is required		
Security updates	Automatic											
Access control	Password - user generated - user changeable, multi-factor authentication, single user account is required											
 Data Practices	<p>Sensor data collection</p> <p>Sensor type</p> <p>Purpose</p> <p>Data stored on device</p> <p>Data stored on cloud</p> <p>Shared with</p> <p>Sold to</p> <p>Other collected data</p> <p>Privacy policy</p>	 Visual Camera Providing device functions No device storage Not disclosed Third parties, government Not sold	 Audio Microphone Providing device functions No device storage Not disclosed Not disclosed Not sold	 Physiological	 Location							
 More Information	<p>Detailed Security & Privacy Label: www.iotsecurityprivacy.org/labels</p>											

Figure 8.3: Most recent version of the primary layer of the label (as of May 2020) that we generated for Ring Doorbell based on the publicly available information. We found the presented information to the best of our ability and the label information has not been verified by the manufacturer.

Security & Privacy Details



Ring Video Doorbell 2
Firmware version: Not disclosed
The device was manufactured in: United States

 Security Mechanisms	Security updates	Automatic +		
	Access control	Password - user generated - user changeable, multi-factor authentication, single user account is required +		
	Security oversight	Not disclosed		
	Ports and protocols	https://support.ring.com/hc/en-us/articles/205385394-The-Protocols-and-Ports-Used-by-Ring-Devices		
	Hardware safety	Not disclosed		
	Software safety	https://shop.ring.com/pages/privacy		
	Personal safety	Not disclosed		
	Vulnerability disclosure and management	Not disclosed		
	Software and hardware composition list	Not disclosed		
	Encryption and key management	https://shop.ring.com/pages/privacy		

 Data Practices	Sensor data collection	Visual	Audio	Movement
	Sensor type	Camera +	Microphone +	Passive infrared motion detectors +
	Collection frequency	Continuous - option to opt out +	Continuous - option to opt out +	Continuous - option to opt out +
	Purpose	Providing device functions +	Providing device functions +	Providing device functions +
	Data stored on device	No device storage +	No device storage +	No device storage +
	Local data retention time	No device storage +	No device storage +	No device storage +
	Data stored on cloud	Not disclosed +	Not disclosed +	Not disclosed +
	Cloud data retention time	60 days +	Not disclosed +	60 days +
	Shared with	Third parties, government +	Not disclosed +	Not disclosed +
	Sharing frequency	Not disclosed, when required by law +	Not disclosed +	Not disclosed +
Sold to	Not sold +	Not sold +	Not sold +	
Other collected data	Contact info, account info, payment info, app geolocation info, product setup info, device tech info, user interaction +			
Data linkage	Not disclosed +			
What could be inferred from user's data	Not disclosed +			
Special data handling practices for children	No +			
In compliance with	GDPR +			
Privacy policy	https://shop.ring.com/pages/privacy			

 More Information	Call Casa with your questions at	1 800 656 1918 +
	Functionality with no internet	Not disclosed +
	Functionality with no data processing	Not disclosed +
	Physical actuations and triggers	Not disclosed +
	Compatible platforms	Amazon Alexa +

Figure 8.4: Most recent version of the secondary layer of the label (as of May 2020) that we generated for Ring Doorbell based on the publicly available information. We found the presented information to the best of our ability and the label information has not been verified by the manufacturer.

8.5.2 Actual Behavior vs. Stated Behavior

Another limitation of our user studies to design the label was that their inputs were self-reported responses. Stated behavior could be different from actual behavior [8]. Therefore, we believe the effectiveness of the label should be evaluated in a realistic setting.

As previously mentioned, various personal and socio-demographic factors could influence the success of labeling. Therefore, future researchers should consider factors such as age, gender, income, level of education, and consumers' motivation when testing the effectiveness of the label in realistic purchase settings.

8.5.3 Monetary Valuation of the Label

One challenge in designing a realistic purchase setting to test the label is incentivizing the IoT manufacturers to label their products in the market for the study. To that end, future work can conduct an incentive-compatible study to specify how much of a premium consumers are willing to pay to have information about the privacy and security practices of their devices. Knowing the amount of premium can help with encouraging manufacturers to voluntarily adopt the labels.

8.5.4 Labels from the Systems Perspective

In this thesis, we described the process of designing a usable and informative label for an IoT device. However, in the era of ambient computing and with the advent of advanced wireless communication standards, such as 5G and beyond, we expect more and more smart devices to connect to the network and communicate with each other. Similar to Manufacturer Usage Description (MUD), the IoT device can broadcast its privacy and security behaviors to all the other network-connected devices via a machine readable format of our label (e.g., JSON) through such mechanisms as the World Wide Web Consortium (W3C) Web of Things (WoT) Description standard that is currently under development [354, 355]. Although transparency at the network level could enable users and network administrators to more effectively detect anomalies in the network, it is imperative to carefully study the privacy implications being introduced by such transparency. Based on the objective of data collection and data sharing, the network administrator should determine which attributes of the IoT label should be publicly available and which attributes need special authorization to be accessed. For example, it might not be safe for everyone to know about the ports and protocols of a device, as they can use this information to attack the device, and thereby the network.

Another future direction is to look beyond the device level and study the privacy and security practices of a network of inter-connected IoT devices. Our designed IoT label describes the privacy and security practices of a single device. By leveraging the concept of a privacy and security label, future work can explore how a label can describe the privacy and security behavior of IoT devices at a system level as a function of the privacy and security label of each individual device in the network.

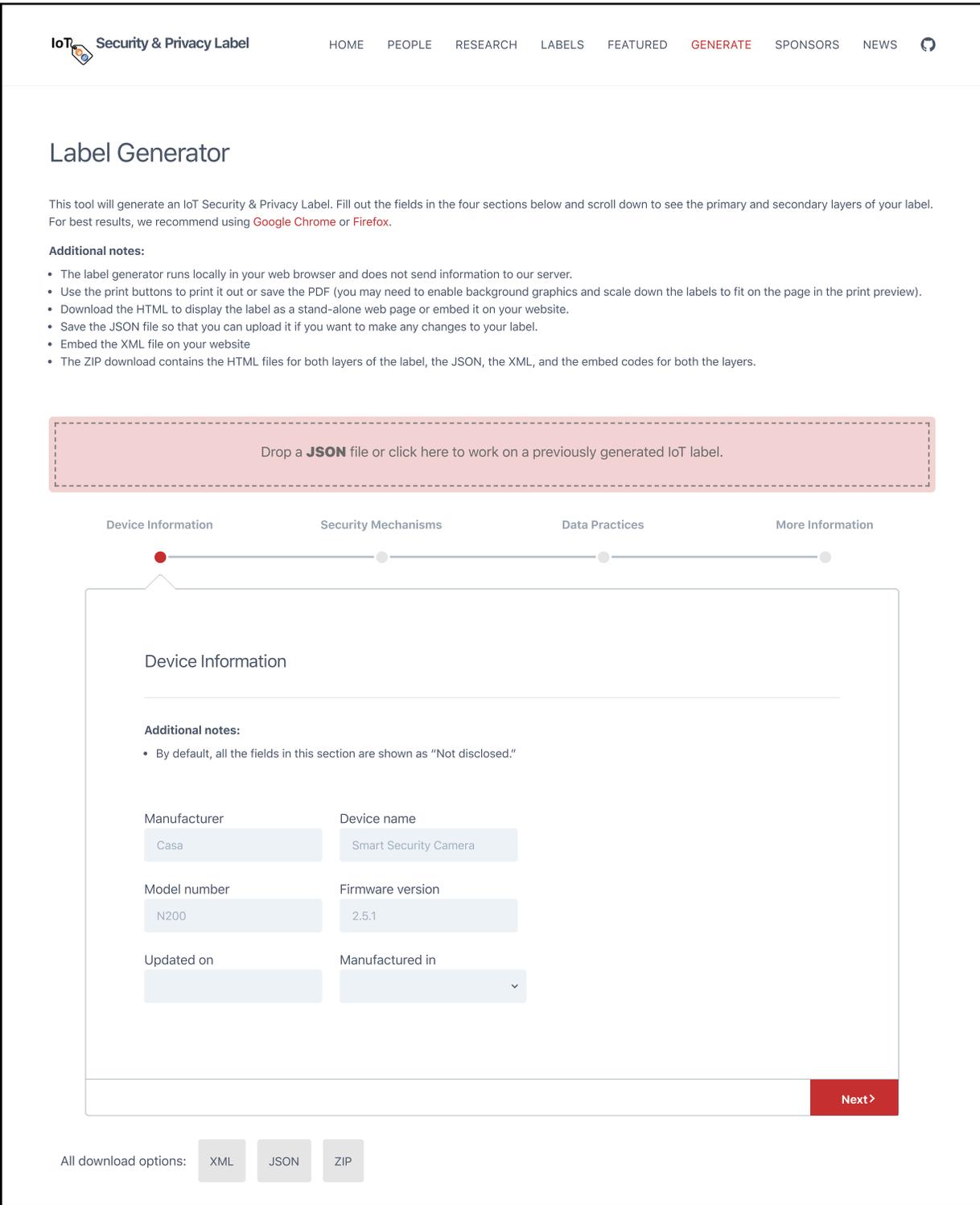


Figure 8.5: The first page of our tool to generate the IoT label.

8.6 Path to Label Adoption

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 an interview study participant in Section 5 mentioned: “I would definitely trust something that had this above something that didn’t.”

Prior work has shown the impact of company size and reputation on consumer trust [173, 337] and purchase behavior [91, 238]. 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 work toward providing protective privacy and security practices to assure their 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 [79]. 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 [124, 138, 139, 140, 141].

In Europe and other countries around the world, enforcement actions could likely be taken by data protection commissioners. The UK government conducted a consultation process from May 1, 2019 to June 5, 2019 to assess three options: mandating retailers to only sell products that have their designed IoT security label, mandating retailers to only sell products that comply with the UK’s previously mentioned “top three” guidelines, and mandating retailers to only sell products that have a label which evidences compliance with all thirteen guidelines from the UK’s Code of Practice for Consumer IoT Security and ETSI TS 103 645. Through this process, they collected 60 formal written responses. The questions were worded in a leading format, mostly starting with “Do you agree,” which might have biased the respondents. Here, we mention some of their findings that could help understand how an IoT label can be adopted and regulated.

Part of the consultation was to ask respondents about their thoughts on government taking power to regulate a security baseline for consumer IoT devices. Based on the responses, the UK government concluded that it is necessary for the government to regulate a baseline for the security of consumer IoT devices. More specifically, the consultation asked whether the aforementioned “top three” guidelines should be considered as the security baseline and from the participants’ feedback, the government decided to move forward with a staging process, starting with mandating the three guidelines as the baseline requirements, while encouraging manufacturers to implement all thirteen guidelines. Moving forward, the UK government will consider all the guidelines included in their Code of Practice.

Another question in the consultation asked about participants’ opinion on the proposed design of the label. A large number of respondents disliked the label design, mainly due to its static nature. Participants reported that a static security label cannot realistically cover the array of current and future IoT technologies and vulnerabilities. We also agree that the rapid pace at which IoT devices receive software and firmware updates could make it a challenge for manufacturers

to keep their static labels up to date. This also means that the adherence of IoT devices' actual behavior to what is on the label is a moving target as features are added or removed, bugs are introduced or fixed, and the firmware gets updated. Our layered design of the label can mitigate this concern by providing manufacturers a space (secondary layer) to update their labels and notify consumers of the updates. The secondary layer is an online-only version of the label that can be accessed by scanning the QR code or typing in the URL provided on the static or the primary layer of the label.

As previously mentioned, one of the main goals of the consultation process was to assess the three options related to the IoT label. Although the government's recommended option was to mandate retailers not to sell consumer IoT products without a security label, respondents expressed a variety of opinions when stating their preferences. A number of participants were in favor of mandating the security label to be on all consumer IoT products sold by the retailers, mostly due to the importance of manufacturers being transparent about their practices. However, there were a number of participants who disagreed with this option. The most common alternative option for this group of participants was to mandate retailers to only sell products that comply with the top three guidelines. Some participants were afraid that by mandating the label instead of the guidelines, the success of the labeling scheme could outweigh the success of minimizing security risks of IoT devices. We agree 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. Therefore, labeling and providing good privacy and security practices are not mutually exclusive. In addition to providing transparency, the label could act as a forcing mechanism to increase market pressure and incentivize manufactures to have better privacy and security practices, hence resulting in more positive labels.

Based on the consultation process, the UK government concluded not to proceed with having a voluntary labeling scheme for now due to the potential challenges for the retailers to validate the manufactures' claims on the label. We believe having a third-party assessment body could address this concern. The UK government argues that having a self-assessment procedure would reduce the manufacturers' cost by empowering them to conduct relevant assessments that are appropriate for their devices.

To conclude, we believe it is important to mandate a security and privacy baseline for IoT devices to ensure a basic level of privacy and security for IoT devices, although there is no consensus on what the baseline should be. In addition to mandating such a baseline, it is imperative to help consumers make more informed decisions when purchasing smart devices. Therefore, having an informative security and privacy label for smart devices is an invaluable undertaking. Our label generator helps manufacturers easily create labels for their smart devices. In the future, having a third-party evaluator could enhance the reliability of the labels and consumers' trust in them.

Bibliography

- [1] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Privacy and human behavior in the age of information. *Science*, 347(6221):509–514, 2015.
- [2] Alessandro Acquisti and Ralph Gross. Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *Proc. PETS*, 2006.
- [3] Alessandro Acquisti and Jens Grossklags. Privacy and rationality in individual decision making. *IEEE security & privacy*, 3(1):26–33, 2005.
- [4] Alessandro Acquisti, Leslie K John, and George Loewenstein. What is privacy worth? *The Journal of Legal Studies*, 42(2):249–274, 2013.
- [5] Michael Adler and Erio Ziglio. *Gazing into the oracle: The Delphi method and its application to social policy and public health*. Jessica Kingsley Publishers, 1996.
- [6] Yuvraj Agarwal and Malcolm Hall. Protectmyprivacy: detecting and mitigating privacy leaks on ios devices using crowdsourcing. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pages 97–110, 2013.
- [7] Herman Aguinis and Kyle J Bradley. Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, 17(4):351–371, 2014.
- [8] Icek Ajzen, Thomas C Brown, and Franklin Carvajal. Explaining the discrepancy between intentions and actions: The case of hypothetical bias in contingent valuation. *Personality and social psychology bulletin*, 30(9):1108–1121, 2004.
- [9] Icek. Ajzen and Martin. Fishbein. *Understanding attitudes and predicting social behavior / Icek Ajzen, Martin Fishbein*. Prentice-Hall Englewood Cliffs, N.J, 1980.
- [10] Angeliki Aktypi, Jason R.C. Nurse, and Michael Goldsmith. Unwinding ariadne’s identity thread: Privacy risks with fitness trackers and online social networks. In *Proceedings of the 2017 on Multimedia Privacy and Security, MPS ’17*, page 1–11, New York, NY, USA, 2017. Association for Computing Machinery.
- [11] Vernon L Allen. Situational factors in conformity¹. In *Advances in experimental social psychology*, volume 2, pages 133–175. Elsevier, 1965.
- [12] Vernon L Allen and John M Levine. Social support and conformity: The role of independent assessment of reality. *Journal of Experimental Social Psychology*, 7(1):48–58, 1971.

- [13] Nick Allum. An empirical test of competing theories of hazard-related trust: The case of gm food. *Risk Analysis: An International Journal*, 27(4):935–946, 2007.
- [14] Omar Alrawi, Chaz Lever, Manos Antonakakis, and Fabian Monrose. SoK: Security evaluation of home-based IoT deployments. In *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, 2019.
- [15] Hamza Alshenqeeti. Interviewing as a data collection method: A critical review. *English Linguistics Research*, 3(1):39, 2014.
- [16] B Charles Ames and James D Hlavacek. *Managerial marketing for industrial firms*. Random House, Business Division, 1984.
- [17] Alan R Andreasen. A taxonomy of consumer satisfaction/dissatisfaction measures. *Journal of Consumer Affairs*, 11(2):11–24, 1977.
- [18] J Craig Andrews, Richard G Netemeyer, and Scot Burton. Consumer generalization of nutrient content claims in advertising. *Journal of marketing*, 62(4):62–75, 1998.
- [19] Manos Antonakakis, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J Alex Halderman, Luca Invernizzi, Michalis Kallitsis, et al. Understanding the Mirai botnet. In *26th USENIX Security Symposium (USENIX Security 17)*, pages 1093–1110, 2017.
- [20] Noah Apthorpe, Dillon Reisman, Srikanth Sundaresan, Arvind Narayanan, and Nick Feamster. Spying on the smart home: Privacy attacks and defenses on encrypted IoT traffic. *arXiv preprint arXiv:1708.05044*, 2017.
- [21] Orlando Arias, Jacob Wurm, Khoa Hoang, and Yier Jin. Privacy and security in internet of things and wearable devices. *IEEE Transactions on Multi-Scale Computing Systems*, 1(2):99–109, 2015.
- [22] Solomon E Asch. Studies of independence and conformity: I. a minority of one against a unanimous majority. *Psychological monographs: General and applied*, 70(9):1, 1956.
- [23] Nor Hazlin Nor Asshidin, Nurazariah Abidin, and Hafizzah Bashira Borhan. Perceived quality and emotional value that influence consumer’s purchase intention towards american and local products. *Procedia Economics and Finance*, 35:639–643, 2016.
- [24] Charles R Atherton. Group techniques for program planning: A guide to nominal group and Delphi processes. by André L. Delbecq, Andrew H. Van de Ven, and David H. Gustafson. Glenview, Ill.: Scott, Foresman & Co., 1976.
- [25] Susan Athey, Christian Catalini, and Catherine Tucker. The digital privacy paradox: Small money, small costs, small talk. Technical report, National Bureau of Economic Research, 2017.
- [26] Christiane Atzmüller and Peter M. Steiner. Experimental vignette studies in survey research. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6(3):128–138, 2010.
- [27] Kristen Backor, Saar Golde, and Norman Nie. Estimating survey fatigue in time use study. In *international association for time use research conference*. Washington, DC. Citeseer, 2007.

- [28] Gökhan Bal, Kai Rannenberg, and Jason I Hong. Styx: Privacy risk communication for the android smartphone platform based on apps' data-access behavior patterns. *Computers & Security*, 53:187–202, 2015.
- [29] Rebecca Balebako, Cristian Bravo-Lillo, and Lorrie Faith Cranor. Is notice enough: Mitigating the risks of smartphone data sharing. *ISJLP*, 11:279, 2015.
- [30] Rebecca Balebako, Jaeyeon Jung, Wei Lu, Lorrie Faith Cranor, and Carolyn Nguyen. Little brothers watching you: Raising awareness of data leaks on smartphones. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, page 12. ACM, 2013.
- [31] Martina Balestra, Orit Shaer, Johanna Okerlund, Madeleine Ball, and Oded Nov. The effect of exposure to social annotation on online informed consent beliefs and behavior. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 900–912, New York, NY, USA, 2016. ACM.
- [32] Adam Bannister. Iot security: Singapore launches labeling scheme for wifi routers and home hubs. <https://portswigger.net/daily-swig/iot-security-singapore-launches-labeling-scheme-for-wifi-routers-and-home-hubs>.
- [33] Mario Ballano Barcena and Candid Wueest. Insecurity in the internet of things. *Security Response, Symantec*, 2015.
- [34] Susan B Barnes. A privacy paradox: Social networking in the united states. *First Monday*, 11(9), 2006.
- [35] Jonathan Baron. *Thinking and deciding*. Cambridge University Press, 2000.
- [36] Robert S Baron, Joseph A Vandello, and Bethany Brunzman. The forgotten variable in conformity research: Impact of task importance on social influence. *Journal of personality and social psychology*, 71(5):915, 1996.
- [37] Christine Barter and Emma Renold. The use of vignettes in qualitative research. *Social research update*, 25(9):1–6, 1999.
- [38] Debjane Barua, Judy Kay, and Cécile Paris. Viewing and Controlling Personal Sensor Data: What Do Users Want? In Shlomo Berkovsky and Jill Freyne, editors, *Persuasive Technology*, number 7822 in Lecture Notes in Computer Science, pages 15–26. Springer Berlin Heidelberg, April 2013. DOI: 10.1007/978-3-642-37157-8_4.
- [39] Debjane Barua, Judy Kay, and Cécile Paris. Viewing and controlling personal sensor data: what do users want? In *International Conference on Persuasive Technology*, pages 15–26. Springer, 2013.
- [40] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015.
- [41] Xavier Bellekens, Andrew Hamilton, Preetila Seem, Kamila Nieradzinska, Quentin Franssen, and Amar Seem. Pervasive ehealth services a security and privacy risk awareness survey. In *2016 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (CyberSA)*, pages 1–4. IEEE, 2016.

- [42] Michael Benisch, Patrick Gage Kelley, Norman Sadeh, and Lorrie Faith Cranor. Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs. *Personal and Ubiquitous Computing*, 15(7):679–694, 2011.
- [43] Andrew Besmer, Jason Watson, and Heather Richter Lipford. The impact of social navigation on privacy policy configuration. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, SOUPS '10, pages 7:1–7:10, New York, NY, USA, 2010. ACM.
- [44] Pankaj Bhaskar and Sheikh Iqbal Ahamed. Privacy in pervasive computing and open issues. In *Proceedings of the Second International Conference on Availability, Reliability and Security, ARES 2007, The International Dependability Conference - Bridging Theory and Practice*, pages 147–154, 2007.
- [45] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5):992–1026, 1992.
- [46] Igor Bilogrevic and Martin Ortlieb. “If You Put All The Pieces Together...”: Attitudes Towards Data Combination and Sharing Across Services and Companies. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, 2016.
- [47] RD Blackwell, PW Miniard, and JF Engell. Consumer behaviour, 10th international student ed. *Thomson South-Western, Mason, OH*, 2006.
- [48] J. M. Blythe and S. D. Johnson. The consumer security index for IoT: A protocol for developing an index to improve consumer decision making and to incentivize greater security provision in IoT devices. In *Living in the Internet of Things: Cybersecurity of the IoT - 2018*, pages 1–7, March 2018.
- [49] John M Blythe, Nissy Sombatruang, and Shane D Johnson. What security features and crime prevention advice is communicated in consumer IoT device manuals and support pages? *Journal of Cybersecurity*, 5(1), 06 2019. tyz005.
- [50] Bram Bonné, Sai Teja Peddinti, Igor Bilogrevic, and Nina Taft. Exploring decision making with Android’s runtime permission dialogs using in-context surveys. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 195–210, 2017.
- [51] Ann Bostrom, Cynthia J Atman, Baruch Fischhoff, and M Granger Morgan. Evaluating risk communications: completing and correcting mental models of hazardous processes, part ii. *Risk analysis*, 14(5):789–798, 1994.
- [52] Nellie Bowles. Thermostats, locks and lights: Digital tools of domestic abuse. <https://www.nytimes.com/2018/06/23/technology/smart-home-devices-domestic-abuse.html>, June 2018.
- [53] Laura Brandimarte, Alessandro Acquisti, and George Loewenstein. Misplaced confidences: Privacy and the control paradox. *Social Psychological and Personality Science*, 4(3):340–347, 2013.
- [54] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.

- [55] Glynis M Breakwell. Risk communication: factors affecting impact. *British medical bulletin*, 56(1):110–120, 2000.
- [56] British Standards Institution. Bsi launches kitemark for internet of things devices. <https://www.bsigroup.com/en-GB/about-bsi/media-centre/press-releases/2018/may/bsi-launches-kitemark-for-internet-of-things-devices/>, March 2018.
- [57] Thomas C Brown and Paul Slovic. Effects of context on economic measures of value. *Amenity resource valuation: Integrating economics with other disciplines*, pages 23–30, 1988.
- [58] Jozef Bucko, Lukáš Kakalejčík, and Martina Ferencová. Online shopping: Factors that affect consumer purchasing behaviour. *Cogent Business & Management*, 5(1):1535751, 2018.
- [59] Michael Buhrmester, Tracy Kwang, and Samuel D Gosling. Amazon’s mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on psychological science*, 6(1):3–5, 2011.
- [60] Samuel Burke. Google admits its new smart speaker was eavesdropping on users. <https://money.cnn.com/2017/10/11/technology/google-home-mini-security-flaw/>, 2017.
- [61] Sandra J Burke, Sandra J Milberg, and Wendy W Moe. Displaying common but previously neglected health claims on product labels: understanding competitive advantages, deception, and education. *Journal of Public Policy & Marketing*, 16(2):242–255, 1997.
- [62] Ronald S Burt. Structural holes versus network closure as social capital. In *Social capital*, pages 31–56. Routledge, 2017.
- [63] Simon Byers, Lorrie Faith Cranor, Dave Kormann, and Patrick McDaniel. Searching for privacy: Design and implementation of a p3p-enabled search engine. In *International Workshop on Privacy Enhancing Technologies*, pages 314–328. Springer, 2004.
- [64] California Energy Commission. Power content label (pcl). http://www.energy.ca.gov/pcl/power_content_label.html, 2009.
- [65] Jen Caltrider. 10 fascinating things we learned when we asked the world “how connected are you?”. <https://goo.gl/92JDfQ>, November 2017.
- [66] Scott Camazine. *Self-organization in biological systems*. Princeton University Press, 2003.
- [67] Richard Chirgwin. Australia’s IoT security rating might work, if done right. https://www.theregister.co.uk/2017/10/17/iot_security_rating_it_can_work_if_done_right/, October 2017.
- [68] Robert B Cialdini and Noah J Goldstein. Social influence: Compliance and conformity. *Annu. Rev. Psychol.*, 55:591–621, 2004.
- [69] Cisco. Building trust and value in the data exchange between people, things and providers. https://www.jasper.com/sites/default/files/cisco_iiot_survey_-_the_value_trust_paradox_final-1_2.pdf, December 2017.

- [70] Robert T Clemen and Robert L Winkler. Combining economic forecasts. *Journal of Business & Economic Statistics*, 4(1):39–46, 1986.
- [71] Samuel W Cochran. The Delphi method: Formulating and refining group judgements. *Journal of Human Sciences*, 2(2):111–117, 1983.
- [72] Catherine A Cole and Siva K Balasubramanian. Age differences in consumers’ search for information: Public policy implications. *Journal of Consumer Research*, 20(1):157–169, 1993.
- [73] Janet Fagan Coleman, Robert R Blake, and Jane Srygley Mouton. Task difficulty and conformity pressures. *The Journal of Abnormal and Social Psychology*, 57(1):120, 1958.
- [74] Jessica Colnago, Yuanyuan Feng, Tharangini Palanivel, Sarah Pearman, Megan Ung, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. Informing the design of a personalized privacy assistant for the internet of things. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 2020.
- [75] Abigail R Colson and Roger M Cooke. Expert elicitation: using the classical model to validate experts’ judgments. *Review of Environmental Economics and Policy*, 12(1):113–132, 2018.
- [76] The Federal Trade Commission. Internet of Things: Privacy & Security in a Connected World. Technical report, Federal Trade Commission, 2015. Accessed Mar. 2017.
- [77] Iain D Couzin, Christos C Ioannou, Güven Demirel, Thilo Gross, Colin J Torney, Andrew Hartnett, Larissa Conrard, Simon A Levin, and Naomi E Leonard. Uninformed individuals promote democratic consensus in animal groups. *science*, 334(6062):1578–1580, 2011.
- [78] Vincent T Covello, W Gary Flamm, Joseph V Rodricks, and Robert G Tardiff. *The analysis of actual versus perceived risks*, volume 1. Springer Science & Business Media, 2012.
- [79] Lorrie Faith Cranor. Necessary but not sufficient: Standardized mechanisms for privacy notice and choice. *J. on Telecomm. & High Tech. L.*, 10:273, 2012.
- [80] Lorrie Faith Cranor, Pedro Giovanni Leon, and Blase Ur. A large-scale evaluation of us financial institutions’ standardized privacy notices. *ACM Transactions on the Web (TWEB)*, 10(3):17, 2016.
- [81] Lorrie Faith Cranor, Joseph Reagle, and Mark S Ackerman. Beyond concern: Understanding net users’ attitudes about online privacy. *The Internet upheaval: raising questions, seeking answers in communications policy*, pages 47–70, 2000.
- [82] Matthew A Cronin and Laurie R Weingart. Representational gaps, information processing, and conflict in functionally diverse teams. *Academy of Management Review*, 32(3):761–773, 2007.
- [83] Ang Cui and Salvatore J Stolfo. A quantitative analysis of the insecurity of embedded network devices: results of a wide-area scan. In *Proceedings of the 26th Annual Computer Security Applications Conference*, pages 97–106. ACM, 2010.
- [84] George Cvetkovich and Patricia L Winter. Trust and social representations of the management of threatened and endangered species. *Environment and Behavior*, 35(2):286–307, 2003.

- [85] Cyber Independent Testing Lab. How we evaluate.
- [86] Frederick R Cyphert and Walter L Gant. The Delphi technique: A case study. *Phi Delta Kappan*, 52(5):272–273, 1971.
- [87] Anne Louise Dailey and James C Holmberg. Delphi-a catalytic strategy for motivating curriculum revision by faculty. *Community/Junior College Quarterly of Research and Practice*, 14(2):129–136, 1990.
- [88] Karl Dake. Orienting dispositions in the perception of risk: An analysis of contemporary worldviews and cultural biases. *Journal of cross-cultural psychology*, 22(1):61–82, 1991.
- [89] Karl Manning Dake. Technology on trial : orienting dispositions toward environmental and health hazards, 1990. Authorized facsimile from UMI Dissertation Services.
- [90] Norman Dalkey and Olaf Helmer. An experimental application of the Delphi method to the use of experts. *Management science*, 9(3):458–467, 1963.
- [91] Michael R Darby and Edi Karni. Free competition and the optimal amount of fraud. *The Journal of law and economics*, 16(1):67–88, 1973.
- [92] Anupam Das, Martin Degeling, Daniel Smullen, and Norman Sadeh. *Personalized Privacy Assistants for the Internet of Things*, 2018.
- [93] Sauvik Das, Adam D.I. Kramer, Laura A. Dabbish, and Jason I. Hong. The role of social influence in security feature adoption. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW '15*, pages 1416–1426, New York, NY, USA, 2015. ACM.
- [94] Thomas H Davenport. *Analytics in Healthcare and the Life Sciences: Strategies, Implementation Methods, and Best Practices*. Pearson Education, 2013.
- [95] Lynn E Davis, Arthur Melmed, and Richard Krop. *Individual preparedness and response to chemical, radiological, nuclear, and biological terrorist attacks*. Rand Corporation, 2003.
- [96] Matt Day, Giles Turner, and Natalia Drozdziak. Amazon workers are listening to what you tell alexa. <https://www.bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alex-a-global-team-reviews-audio>, 2019.
- [97] Tamara Denning, Tadayoshi Kohno, and Henry M Levy. Computer security and the modern home. *Communications of the ACM*, 56(1):94–103, 2013.
- [98] Tamara Denning, Cynthia Matuszek, Karl Koscher, Joshua R Smith, and Tadayoshi Kohno. A spotlight on security and privacy risks with future household robots: attacks and lessons. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 105–114. ACM, 2009.

- [99] Department for Digital, Culture, Media & Sport. Consultation on the government's regulatory proposals regarding consumer internet of things (IoT) security. <https://www.gov.uk/government/consultations/consultation-on-regulatory-proposals-on-consumer-iot-security/consultation-on-the-governments-regulatory-proposals-regarding-consumer-internet-of-things-iot-security>, May 2019.
- [100] Department for Digital, Culture, Media and Sport. Code of practice for consumer IoT security. <https://www.gov.uk/government/publications/code-of-practice-for-consumer-iot-security/code-of-practice-for-consumer-iot-security>.
- [101] Department for Digital, Culture, Media and Sport. Statement of intent regarding the security of the internet of things. <https://www.gov.uk/government/publications/five-country-ministerial-communicue/statement-of-intent-regarding-the-security-of-the-internet-of-things>.
- [102] Department for Digital, Culture, Media and Sport. Secure by design: Improving the cyber security of consumer internet of things report. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/686089/Secure_by_Design_Report.pdf, March 2018.
- [103] Department for Digital, Culture, Media and Sport. Mandating security requirements for consumer IoT products. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/798722/Secure_by_Design_Consultation_Stage_Regulatory_Impact_Assessment.pdf, May 2019.
- [104] Department for Digital, Culture, Media and Sport. Plans announced to introduce new laws for internet connected devices. <https://www.gov.uk/government/news/plans-announced-to-introduce-new-laws-for-internet-connected-devices>, May 2019.
- [105] Morton Deutsch and Harold B Gerard. A study of normative and informational social influences upon individual judgment. *The journal of abnormal and social psychology*, 51(3):629, 1955.
- [106] Paul DiGioia and Paul Dourish. Social navigation as a model for usable security. In *Proceedings of the 2005 symposium on Usable privacy and security*, pages 101–108. ACM, 2005.
- [107] Mary Douglas and Aaron Wildavsky. *Risk and culture: An essay on the selection of technological and environmental dangers*. Univ of California Press, 1983.
- [108] Nora A. Draper. From Privacy Pragmatist to Privacy Resigned: Challenging Narratives of Rational Choice in Digital Privacy Debates: Challenging Rational Choice in Digital Privacy Debates. *Policy & Internet*, 9(2):232–251, June 2017.

- [109] Janna Lynn Dupree, Richard Devries, Daniel M. Berry, and Edward Lank. Privacy Personas: Clustering Users via Attitudes and Behaviors Toward Security Practices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 5228–5239, New York, NY, USA, 2016. ACM.
- [110] Catherine Dwyer, Starr Roxanne Hiltz, and Katia Passerini. Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. In *Proc. AMCIS*, 2007.
- [111] E. Lear, R. Droms, and D. Romascanu. Manufacturer Usage Description Specification. Internet-Draft draft-ietf-opsawg-mud-04, IETF Network Working Group, February 2017.
- [112] Timothy C Earle and George Cvetkovich. Social trust and culture in risk management. In *Social trust and the management of risk*, pages 23–35. Routledge, 2013.
- [113] Serge Egelman, Janice Tsai, Lorrie Faith Cranor, and Alessandro Acquisti. Timing is everything?: The effects of timing and placement of online privacy indicators. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 319–328. ACM, 2009.
- [114] Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Cranor, and Hanan Hibshi. Ask the experts: What should be on an IoT privacy and security label? *arXiv preprint arXiv:2002.04631*, 2020.
- [115] Pardis Emami-Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Cranor, and Norman Sadeh. Privacy expectations and preferences in an IoT world. In *SOUPS '17: Proceedings of the 13th Symposium on Usable Privacy and Security*, July 2017.
- [116] Pardis Emami Naeini, Martin Degeling, Lujo Bauer, Richard Chow, Lorrie Faith Cranor, Mohammad Reza Haghighat, and Heather Patterson. The influence of friends and experts on privacy decision making in IoT scenarios. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–26, 2018.
- [117] Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. Exploring how privacy and security factor into IoT device purchase behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 534. ACM, 2019.
- [118] ENISA. Baseline security recommendations for IoT. <https://www.enisa.europa.eu/publications/baseline-security-recommendations-for-iot>.
- [119] EPA. Learn about the label. <https://www.fueleconomy.gov/feg/Find.do?action=bt1>, 2012.
- [120] ETSI. Cyber security for consumer internet of things. https://www.etsi.org/deliver/etsi_ts/103600_103699/103645/01.01.01_60/ts_103645v010101p.pdf.

- [121] European Commission. Proposal for a regulation of the European Parliament and of the Council on ENISA, the "EU cybersecurity agency", and repealing regulation (EU) 526/2013, and on information and communication technology cybersecurity certification ("Cybersecurity Act"). <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52017PC0477&rid=1>, 2017.
- [122] Fariborz Farahmand and Eugene H Spafford. Understanding insiders: An analysis of risk-taking behavior. *Information systems frontiers*, 15(1):5–15, 2013.
- [123] FDA. Nutrition facts label better informs your food choices. <https://www.fda.gov/ForConsumers/ConsumerUpdates/ucm387114.htm>, August 2016.
- [124] Federal Trade Commission. Comment to national telecommunications and information administration. <https://www.ftc.gov/policy/advocacy/advocacy-filings/2017/06/ftc-comment-national-telecommunications-information>, June 2017.
- [125] Leon Festinger. A theory of social comparison processes. *Human relations*, 7(2):117–140, 1954.
- [126] Janet Finch. The vignette technique in survey research. *Sociology*, pages 105–114, 1987.
- [127] Janet Finch. The vignette technique in survey research. *Sociology*, 21(1):105–114, 1987.
- [128] Finnish Transport and Communication Agency. Finland becomes the first european country to certify safe smart devices – new cybersecurity label helps consumers buy safer products. <https://www.traficom.fi/en/news/finland-becomes-first-european-country-certify-safe-smart-devices-new-cybersecurity-label>.
- [129] Baruch Fischhoff, Ann Bostrom, and Marilyn Jacobs Quadrel. Risk perception and communication. *Annual review of public health*, 14(1):183–203, 1993.
- [130] Baruch Fischhoff, Paul Slovic, Sarah Lichtenstein, Stephen Read, and Barbara Combs. How safe is safe enough? a psychometric study of attitudes towards technological risks and benefits. *Policy sciences*, 9(2):127–152, 1978.
- [131] Joseph L Fleiss, Bruce Levin, and Myunghee Cho Paik. *Statistical methods for rates and proportions*. John Wiley & Sons, 2013.
- [132] Nuno Fortes, Paulo Rita, and Margherita Pagani. The effects of privacy concerns, perceived risk and trust on online purchasing behaviour. *International Journal of Internet Marketing and Advertising*, 11(4):307–329, 2017.
- [133] OWASP Foundation. IoT security guidance. https://www.owasp.org/index.php/IoT_Security_Guidance, February 2017.
- [134] Cynthia Franklin and Michelle Ballan. Reliability and validity in qualitative research. *The handbook of social work research methods*, 4:273–292, 2001.
- [135] Batya Friedman, David Hurley, Daniel C Howe, Helen Nissenbaum, and Edward Feltten. Users' conceptions of risks and harms on the web: a comparative study. In *CHI'02 extended abstracts on Human factors in computing systems*, pages 614–615, 2002.

- [136] FTC. Shopping for light bulbs. <https://www.consumer.ftc.gov/articles/0164-shopping-light-bulbs>, april 2011.
- [137] FTC. Shopping for home appliances? use the energyguide label. <https://www.consumer.ftc.gov/articles/0072-shopping-home-appliances-use-energyguide-label>, January 2015.
- [138] FTC. Acdi group llc. <https://www.ftc.gov/enforcement/cases-proceedings/162-3103/acdi-group-llc>, June 2017.
- [139] FTC. Blue global and christopher kay. <https://www.ftc.gov/enforcement/cases-proceedings/152-3225/blue-global-christopher-kay>, July 2017.
- [140] FTC. Lenovo, inc. <https://www.ftc.gov/enforcement/cases-proceedings/152-3134/lenovo-inc>, September 2017.
- [141] FTC. Uber technologies, inc. <https://www.ftc.gov/enforcement/cases-proceedings/152-3054/uber-technologies-inc>, April 2018.
- [142] Vaibhav Garg, Kevin Benton, and L Jean Camp. The privacy paradox: a facebook case study. In *2014 TPRC conference paper*, 2014.
- [143] Vaibhav Garg and Jean Camp. End user perception of online risk under uncertainty. In *2012 45th Hawaii International Conference on System Sciences*, pages 3278–3287. IEEE, 2012.
- [144] Gartner. Internet of things endpoint spending worldwide by category from 2014 to 2020 (in billion U.S. dollars). <https://www.statista.com/statistics/485252/iot-endpoint-spending-by-category-worldwide/>, August 2018.
- [145] David Gefen. E-commerce: the role of familiarity and trust. *Omega*, 28(6):725–737, 2000.
- [146] David Gefen, Izak Benbasat, and Paula Pavlou. A research agenda for trust in online environments. *Journal of Management Information Systems*, 24(4):275–286, 2008.
- [147] Nina Gerber, Benjamin Reinheimer, and Melanie Volkamer. Investigating people’s privacy risk perception. *Proceedings on Privacy Enhancing Technologies*, 2019(3):267–288, 2019.
- [148] Francesca Gino and Don A Moore. Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making*, 20(1):21–35, 2007.
- [149] Global System for Mobile Communications Association (GSMA). GSMA IoT security guidelines for endpoint ecosystems,. <https://www.gsma.com/iot/wp-content/uploads/2016/02/CLP.13-v1.0.pdf>.
- [150] Jeremy Goecks, W Keith Edwards, and Elizabeth D Mynatt. Challenges in supporting end-user privacy and security management with social navigation. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, page 5. ACM, 2009.

- [151] Jeremy Goecks and Elizabeth D Mynatt. Supporting privacy management via community experience and expertise. In *Communities and Technologies 2005*, pages 397–417. Springer, 2005.
- [152] David Goldman. Your Samsung TV is eavesdropping on your private conversations, February 2015.
- [153] Joseph K Goodman, Cynthia E Cryder, and Amar Cheema. Data collection in a flat world: The strengths and weaknesses of mechanical turk samples. *Journal of Behavioral Decision Making*, 26(3):213–224, 2013.
- [154] Jorge Granjal, Edmundo Monteiro, and Jorge Sá Silva. Security for the internet of things: a survey of existing protocols and open research issues. *IEEE Communications Surveys & Tutorials*, 17(3):1294–1312, 2015.
- [155] Andy Greenberg. The reaper IoT botnet has already infected a million networks. <https://www.wired.com/story/reaper-iot-botnet-infected-million-networks/>, September 2017.
- [156] Richard F Griffiths. *Dealing with Risk: The planning, management and acceptability of technological risk*. Manchester University Press, 1981.
- [157] Reeyaz Hamirani. New study: The 2015 state of consumer privacy and personalization. <https://www.gigya.com/blog/new-study-the-2015-state-of-consumer-privacy-personalization/>, July 2018.
- [158] Marian Harbach, Sascha Fahl, and Matthew Smith. Who’s afraid of which bad wolf? a survey of it security risk awareness. In *2014 IEEE 27th Computer Security Foundations Symposium*, pages 97–110. IEEE, 2014.
- [159] Harris Interactive. Consumer internet of things security labelling survey research findings. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/798543/Harris_Interactive_Consumer_IoT_Security_Labelling_Survey_Report.pdf.
- [160] Nigel Harvey and Ilan Fischer. Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes*, 70(2):117–133, 1997.
- [161] Chip Heath, Richard P Larrick, and Joshua Klayman. Cognitive repairs: How organizational practices can compensate for individual shortcomings. In *Review of Organizational Behavior*. Citeseer, 1998.
- [162] Paul Hewson. Statistical rethinking: a bayesian course with examples in r and stan r. mcelreath, 2015 boca raton chapman and hall–crc 470 pp.,£ 60.99 isbn 978-1-482-25344-3. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(4):1131–1132, 2016.
- [163] Hanan Hibshi. *Composite Security Requirements in the Presence of Uncertainty*. PhD thesis, Carnegie Mellon University, 2018.

- [164] Hanan Hibshi, Travis D. Breaux, Maria Riaz, and Laurie Williams. A grounded analysis of experts' decision-making during security assessments. *Journal of Cybersecurity*, 2(2):147–163, 2016.
- [165] Bert H Hodges, Benjamin R Meagher, Daniel J Norton, Ryan McBain, and Ariane Sroubek. Speaking from ignorance: Not agreeing with others we believe are correct. *Journal of Personality and Social Psychology*, 106(2):218, 2014.
- [166] David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. *Applied logistic regression*, volume 398. John Wiley & Sons, 2013.
- [167] Information Commissioner's Office. Energy efficiency. <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2017/11/blog-the-12-ways-that-christmas-shoppers-can-keep-children-and-data-safe-when-buying-smart-toys-and-devices/>, November 2017.
- [168] Information Commissioner's Office. Right to be informed. <https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-be-informed/>, September 2018.
- [169] Harris Interactive. A survey of consumer privacy attitudes and behaviors. *Rochester, NY*, 47, 2000.
- [170] Internet Society. Internet of things (iot) security policy platform. <https://www.internetsociety.org/wp-content/uploads/2019/11/IoT-Security-Platform-EN.pdf>.
- [171] IoT Security Foundation. IoT security compliance framework. <https://www.iotsecurityfoundation.org/wp-content/uploads/2018/12/IoTSF-IoT-Security-Compliance-Framework-Release-2.0-December-2018.pdf>, 2018.
- [172] ioXt. The ioXt security pledge. <https://static1.squarespace.com/static/5c6dbac1f8135a29c7fbb621/t/5ca695ffee6eb0769f5608d1/1554421249364/ioXt-SecurityPledge-booklet-final.pdf>.
- [173] Sirkka L Jarvenpaa, Noam Tractinsky, and Michael Vitale. Consumer trust in an internet store. *Information technology and management*, 1(1-2):45–71, 2000.
- [174] Arthur Jenness. The role of discussion in changing opinion regarding a matter of fact. *The Journal of Abnormal and Social Psychology*, 27(3):279, 1932.
- [175] Helmut Jungermann and Katrin Fischer. Using expertise and experience for giving and taking advice. *The routines of decision making*, pages 157–173, 2005.
- [176] Joseph B Kadane and Nicole A Lazar. Methods and criteria for model selection. *Journal of the American statistical Association*, 99(465):279–290, 2004.
- [177] Ruogu Kang, Stephanie Brown, Laura Dabbish, and Sara B Kiesler. Privacy attitudes of mechanical turk workers and the us public. In *SOUPS*, pages 37–49, 2014.

- [178] Fahri Karakaya and Nora Ganim Barnes. Impact of online reviews of customer care experience on brand or company selection. *Journal of Consumer Marketing*, 27(5):447–457, 2010.
- [179] Heikki Karjaluoto, Jari Karvonen, Manne Kesti, Timo Koivumäki, Marjukka Manninen, Jukka Pakola, Annu Ristola, and Jari Salo. Factors affecting consumer choice of mobile phones: Two studies from finland. *Journal of Euromarketing*, 14(3):59–82, 2005.
- [180] Surya Mattu Kashmir Hill. The house that spied on me. <https://gizmodo.com/the-house-that-spied-on-me-1822429852>, February 2018.
- [181] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W Reeder. A nutrition label for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, page 4. ACM, 2009.
- [182] Patrick Gage Kelley, Lucian Cesca, Joanna Bresee, and Lorrie Faith Cranor. Standardizing privacy notices: an online study of the nutrition label approach. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, pages 1573–1582. ACM, 2010.
- [183] Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. Privacy as part of the app decision-making process. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3393–3402. ACM, 2013.
- [184] J Patrick Kelly and Richard T Hise. Industrial and consumer goods product managers are different. *Industrial Marketing Management*, 8(4):325–332, 1979.
- [185] Herbert C Kelman. Compliance, identification, and internalization three processes of attitude change. *Journal of conflict resolution*, 2(1):51–60, 1958.
- [186] Jennifer King and Aylin Selcukoglu. Where’s the beep? a case study of user misunderstandings of rfid. In *2011 IEEE International Conference on RFID*, pages 192–199. IEEE, 2011.
- [187] Predrag Klasnja, Sunny Consolvo, Tanzeem Choudhury, Richard Beckwith, and Jeffrey Hightower. Exploring Privacy Concerns about Personal Sensing. In Hideyuki Tokuda, Michael Beigl, Adrian Friday, A. J. Bernheim Brush, and Yoshito Tobe, editors, *Pervasive Computing*, number 5538 in Lecture Notes in Computer Science, pages 176–183. Springer Berlin Heidelberg, May 2009. DOI: 10.1007/978-3-642-01516-8_13.
- [188] Predrag Klasnja, Sunny Consolvo, Tanzeem Choudhury, Richard Beckwith, and Jeffrey Hightower. Exploring privacy concerns about personal sensing. In *International Conference on Pervasive Computing*, pages 176–183. Springer, 2009.
- [189] Pamela Klopp and Maurice MacDonald. Nutrition labels: an exploratory study of consumer reasons for nonuse. *Journal of Consumer Affairs*, 15(2):301–316, 1981.
- [190] B. P. Knijnenburg. Privacy? I Can’t Even! Making a Case for User-Tailored Privacy. *IEEE Security & Privacy*, 15(4):62–67, 2017.
- [191] Anne B Knol, Pauline Slottje, Jeroen P van der Sluijs, and Erik Lebret. The use of expert elicitation in environmental health impact assessment: a seven step procedure. *Environmental Health*, 9(1):19, 2010.

- [192] Joseph A. Konstan and John Riedl. *Collaborative Filtering: Supporting Social Navigation in Large, Crowded Infospaces*, pages 43–82. Springer London, London, 2003.
- [193] Petra M Kuhnert, Tara G Martin, and Shane P Griffiths. A guide to eliciting and using expert knowledge in bayesian ecological models. *Ecology letters*, 13(7):900–914, 2010.
- [194] Ponnurangam Kumaraguru and Lorrie Cranor. Privacy indexes : a survey of Westin’s studies. *Institute for Software Research*, January 2005.
- [195] J. Richard Landis and Gary G. Koch. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1):159, March 1977.
- [196] Veronica Lara. What the internet of things means for consumer privacy. <https://perspectives.eiu.com/technology-innovation/what-internet-things-means-consumer-privacy-0/white-paper/what-internet-things-means-consumer-privacy>, March 2018.
- [197] Richard P Larrick and Jack B Soll. Intuitions about combining opinions: Misappreciation of the averaging principle. *Management science*, 52(1):111–127, 2006.
- [198] Bibb Latané. The psychology of social impact. *American psychologist*, 36(4):343, 1981.
- [199] Leadership in Energy and Environmental Design. Green building leadership is lead.
- [200] Scott Lederer, Jennifer Mankoff, and Anind K. Dey. Who wants to know what when? privacy preference determinants in ubiquitous computing. In *Extended abstracts of the 2003 Conference on Human Factors in Computing Systems, CHI 2003, Ft. Lauderdale, Florida, USA, April 5-10, 2003*, pages 724–725, 2003.
- [201] Scott Lederer, Jennifer Mankoff, and Anind K Dey. Who wants to know what when? privacy preference determinants in ubiquitous computing. In *CHI’03 extended abstracts on Human factors in computing systems*, pages 724–725. ACM, 2003.
- [202] Hosub Lee and Alfred Kobsa. Understanding user privacy in internet of things environments. In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, pages 407–412. IEEE, 2016.
- [203] Hosub Lee and Alfred Kobsa. Privacy preference modeling and prediction in a simulated campuswide IoT environment. In *2017 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 276–285. IEEE, 2017.
- [204] Hosub Lee and Alfred Kobsa. Privacy preference modeling and prediction in a simulated campuswide IoT environment. In *2017 IEEE International Conference on Pervasive Computing and Communications, PerCom 2017, Hawaii, USA, March 13-17, 2017*, pages 276–285, 2017.
- [205] Linda Lee, J Lee, Serge Egelman, and David Wagner. Information disclosure concerns in the age of wearable computing. In *NDSS Workshop on Usable Security (USEC)*, volume 1, 2016.
- [206] Pedro Giovanni Leon, Blase Ur, Yang Wang, Manya Sleeper, Rebecca Balebako, Richard Shay, Lujo Bauer, Mihai Christodorescu, and Lorrie Faith Cranor. What matters to users?: factors that affect users’ willingness to share information with online advertisers. In *Proceedings of the ninth symposium on usable privacy and security*, page 7. ACM, 2013.

- [207] Pedro Giovanni Leon, Blase Ur, Yang Wang, Manya Sleeper, Rebecca Balebako, Richard Shay, Lujo Bauer, Mihai Christodorescu, and Lorrie Faith Cranor. What matters to users?: Factors that affect users' willingness to share information with online advertisers. In *Proceedings of the ninth symposium on usable privacy and security*, page 7. ACM, 2013.
- [208] Rachel Lerman. Google's Nest Hub has a microphone it forgot to mention. <https://www.usnews.com/news/business/articles/2019-02-20/googles-nest-hub-has-a-microphone-it-forgot-to-mention>, 2019.
- [209] Ted Lieu. H.R.4163: Cyber Shield Act of 2017. <https://www.congress.gov/115/bills/hr4163/BILLS-115hr4163ih.pdf>, October 2017.
- [210] Jialiu Lin, Shahriyar Amini, Jason I Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 501–510. ACM, 2012.
- [211] Jialiu Lin, Bin Liu, Norman Sadeh, and Jason I Hong. Modeling users' mobile app privacy preferences: Restoring usability in a sea of permission settings. In *Symposium on Usable Privacy and Security (SOUPS)*, volume 40, 2014.
- [212] Chen Ling, Wonil Hwang, and Gavriel Salvendy. Diversified users' satisfaction with advanced mobile phone features. *Universal Access in the Information Society*, 5(2):239–249, 2006.
- [213] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhiemedi, Shikun Zhang, Norman M. Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Twelfth Symposium on Usable Privacy and Security, SOUPS 2016, Denver, CO, USA, June 22-24, 2016*, pages 27–41, 2016.
- [214] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhiemedi, Shikun Aerin Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. Follow my recommendations: A personalized assistant for mobile app permissions. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, 2016.
- [215] Bin Liu, Jialiu Lin, and Norman Sadeh. Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help? In *Proceedings of the 23rd international conference on World wide web*, pages 201–212. ACM, 2014.
- [216] Steson Lo and Sally Andrews. To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, 6:1171, 2015.
- [217] Jan Lorenz, Heiko Rauhut, Frank Schweitzer, and Dirk Helbing. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, 108(22):9020–9025, 2011.
- [218] Todd Lucas, Sheldon Alexander, Ira J Firestone, and Boris B Baltes. Self-efficacy and independence from social influence: Discovery of an efficacy–difficulty effect. *Social Influence*, 1(1):58–80, 2006.

- [219] Barbara Ludwig. Predicting the future: Have you considered using the Delphi methodology. *Journal of extension*, 35(5):1–4, 1997.
- [220] Zoë Mack and Sarah Sharples. The importance of usability in product choice: A mobile phone case study. *Ergonomics*, 52(12):1514–1528, 2009.
- [221] Diane M Mackie. Systematic and nonsystematic processing of majority and minority persuasive communications. *Journal of Personality and Social Psychology*, 53(1):41, 1987.
- [222] Kathleen M. MacQueen, Eleanor McLellan, Kelly Kay, and Bobby Milstein. Codebook Development for Team-Based Qualitative Analysis. *CAM Journal*, 10(2):31–36, May 1998.
- [223] Kathleen M Macqueen, Eleanor McLellan-Lemal, Kelly Bartholow, and Bobby Milstein. Team-based codebook development: Structure, process, and agreement. *Handbook for team-based qualitative research*, pages 119–135, 2008.
- [224] Anna Madill, Abbie Jordan, and Caroline Shirley. Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies. *British journal of psychology*, 91(1):1–20, 2000.
- [225] Naresh K Malhotra, Sung S Kim, and James Agarwal. Internet users’ information privacy concerns (iuipe): The construct, the scale, and a causal model. *Information systems research*, 15(4):336–355, 2004.
- [226] Albert E Mannes. Are we wise about the wisdom of crowds? the use of group judgments in belief revision. *Management Science*, 55(8):1267–1279, 2009.
- [227] Carsten Maple. Security and privacy in the internet of things. *Journal of Cyber Policy*, 2(2):155–184, 2017.
- [228] Edward Markey. S.2020: Cyber shield act of 2017. <https://www.congress.gov/115/bills/s2020/BILLS-115s2020is.pdf>, October 2017.
- [229] David F Marks and Lucy Yardley. *Research methods for clinical and health psychology*. Sage, 2004.
- [230] Kirsten E Martin. Diminished or just different? a factorial vignette study of privacy as a social contract. *Journal of Business Ethics*, 111(4):519–539, 2012.
- [231] Kirsten E Martin and Helen Nissenbaum. Measuring privacy: An empirical test using context to expose confounding variables. *Columbia Science and Technology Law Review*, 18:176–218, 2016.
- [232] Robin Martin, Antonis Gardikiotis, and Miles Hewstone. Levels of consensus and majority and minority influence. *European Journal of Social Psychology*, 32(5):645–665, 2002.
- [233] Alan D Mathios. Socioeconomic factors, nutrition, and food choices: An analysis of the salad dressing market. *Journal of Public Policy & Marketing*, 15(1):45–54, 1996.

- [234] Arunesh Mathur, Gunes Acar, Michael J Friedman, Elena Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan. Dark patterns at scale: Findings from a crawl of 11k shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–32, 2019.
- [235] Pavlin Mavrodiev, Claudio J Tessone, and Frank Schweitzer. Effects of social influence on the wisdom of crowds. *arXiv preprint arXiv:1204.3463*, 2012.
- [236] Pavlin Mavrodiev, Claudio J Tessone, and Frank Schweitzer. Quantifying the effects of social influence. *Scientific reports*, 3:1360, 2013.
- [237] Alex Mayle, Neda Hajiakhoond Bidoki, Sina Masnadi, Ladislau Boeloeni, and Damla Turgut. Investigating the value of privacy within the internet of things. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pages 1–6. IEEE, 2017.
- [238] M Mazzocchi, AE Lobb, and BW Traill. A strategy for measuring trust in food safety information: A literature review. Technical report, University of Florence Working Paper Series on Trust, 2004.
- [239] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for cscw and hci practice. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), November 2019.
- [240] William McGeeveran. The law of friction. *U. Chi. Legal F.*, page 15, 2013.
- [241] Frank P McKenna. It won't happen to me: Unrealistic optimism or illusion of control? *British Journal of Psychology*, 84(1):39–50, 1993.
- [242] Anneloes Meijnders, Cees Midden, Anna Olofsson, Susanna Öhman, Jörg Matthes, Olha Bondarenko, Jan Gutteling, and Maria Rusanen. The role of similarity cues in the development of trust in sources of information about gm food. *Risk Analysis: An International Journal*, 29(8):1116–1128, 2009.
- [243] Tamir Mendel and Eran Toch. Susceptibility to social influence of privacy behaviors: Peer versus authoritative sources. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '17*, pages 581–593, New York, NY, USA, 2017. ACM.
- [244] Dar Meshi, Guido Biele, Christoph W Korn, and Hauke R Heekeren. How expert advice influences decision making. *PLoS One*, 7(11):e49748, 2012.
- [245] Tomasz Miaskiewicz and Kenneth Kozar. The use of the Delphi method to determine the benefits of the personas method-an approach to systems design. *SIGHCI 2006 Proceedings*, page 7, 2006.
- [246] Benjamin Michéle and Andrew Karpow. Watch and be watched: Compromising all smart tv generations. In *2014 IEEE 11th Consumer Communications and Networking Conference (CCNC)*, pages 351–356. IEEE, 2014.
- [247] Lisa M Soederberg Miller and Diana L Cassidy. Making healthy food choices using nutrition facts panels. the roles of knowledge, motivation, dietary modifications goals, and age. *Appetite*, 59(1):129–139, 2012.

- [248] Lisa M Soederberg Miller and Diana L Cassady. Making healthy food choices using nutrition facts panels. the roles of knowledge, motivation, dietary modifications goals, and age. *Appetite*, 59(1):129–139, 2012.
- [249] Lisa M Soederberg Miller, Tanja N Gibson, and Elizabeth A Applegate. Predictors of nutrition information comprehension in adulthood. *Patient Education and Counseling*, 80(1):107–112, 2010.
- [250] Christine Moorman. The effects of stimulus and consumer characteristics on the utilization of nutrition information. *Journal of Consumer Research*, 17(3):362–374, 1990.
- [251] Christine Moorman, Kristin Diehl, David Brinberg, and Blair Kidwell. Subjective knowledge, search locations, and consumer choice. *Journal of Consumer Research*, 31(3):673–680, 2004.
- [252] A. Morton and M. A. Sasse. Desperately seeking assurances: Segmenting users by their information-seeking preferences. In *2014 Twelfth Annual International Conference on Privacy, Security and Trust*, pages 102–111, July 2014.
- [253] Mehdi Moussaid, Simon Garnier, Guy Theraulaz, and Dirk Helbing. Collective information processing and pattern formation in swarms, flocks, and crowds. *Topics in Cognitive Science*, 1(3):469–497, 2009.
- [254] Mozilla. Shop safe this holiday season. <https://foundation.mozilla.org/en/privacynotincluded/>, 2018.
- [255] John W Murry Jr and James O Hammons. Delphi: A versatile methodology for conducting qualitative research. *The Review of Higher Education*, 18(4):423–436, 1995.
- [256] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorie Faith Cranor, and Norman Sadeh. Privacy expectations and preferences in an IoT world. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 399–412, 2017.
- [257] Kazuya Nakayachi and George Cvetkovich. Public trust in government concerning tobacco control in japan. *Risk Analysis: An International Journal*, 30(1):143–152, 2010.
- [258] Rodolfo M Nayga Jr. Nutrition knowledge, gender, and food label use. *Journal of Consumer Affairs*, 34(1):97–112, 2000.
- [259] N Neil, P Slovic, and PJ Hakkinen. Mapping consumer perceptions of risk. *Washington, DC: Chem. Manufactures Assoc*, 1993.
- [260] Raymond S Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2):175, 1998.
- [261] Svein-Egil Nielsen. Should connected devices carry an IoT security-star rating? <https://internetofthingsagenda.techtarget.com/blog/IoT-Agenda/Should-connected-devices-carry-an-IoT-security-star-rating>, May 2019.
- [262] Helen Nissenbaum. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press, 2009.

- [263] Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, November 2009.
- [264] Rajyalakshmi Nittala. Factors influencing online shopping behavior of urban consumers in india. *International Journal of Online Marketing (IJOM)*, 5(1):38–50, 2015.
- [265] Patricia A Norberg, Daniel R Horne, and David A Horne. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, 41(1):100–126, 2007.
- [266] Geoff Norman. Likert scales, levels of measurement and the “laws” of statistics. *Advances in health sciences education*, 15(5):625–632, 2010.
- [267] NTIA. Communicating iot device security update capability to improve transparency for consumers. https://www.ntia.doc.gov/files/ntia/publications/communicating_iot_security_update_capability_for_consumers_-_jul_2017.pdf, 2017.
- [268] Office of Energy Efficiency & Renewable Energy. Energy star. <https://www.energy.gov/eere/buildings/energy-star>.
- [269] Richard L Oliver and William O Bearden. Crossover effects in the theory of reasoned action: A moderating influence attempt. *Journal of consumer research*, 12(3):324–340, 1985.
- [270] Christi Olson. New report tackles tough questions on voice and AI. <https://about.ads.microsoft.com/en-us/blog/post/april-2019/new-report-tackles-tough-questions-on-voice-and-ai>, 2019.
- [271] Jerry C Olson and Jacob Jacoby. Cue utilization in the quality perception process. *ACR Special Volumes*, 1972.
- [272] Temitope Oluwafemi, Tadayoshi Kohno, Sidhant Gupta, and Shwetak Patel. Experimental security analyses of non-networked compact fluorescent lamps: A case study of home automation security. In *LASER*, pages 13–24, 2013.
- [273] Online Trust Alliance. Ota iot trust framework. <https://www.internetsociety.org/iot/trust-framework/>.
- [274] OWASP. Top IoT vulnerabilities. https://www.owasp.org/index.php/Top_IoT_Vulnerabilities, May 2016.
- [275] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. Running experiments on amazon mechanical turk. *Judgment and Decision making*, 5(5):411–419, 2010.
- [276] Sameer Patil, Xinru Page, and Alfred Kobsa. With a little help from my friends: can social navigation inform interpersonal privacy preferences? In *Proceedings of the 2011 ACM Conference on Computer Supported Cooperative Work, CSCW 2011, Hangzhou, China, March 19-23, 2011*, pages 391–394, 2011.
- [277] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

- [278] Charith Perera, Rajiv Ranjan, Lizhe Wang, Samee Ullah Khan, and Albert Y. Zomaya. Big data privacy in the internet of things era. *IT Professional*, 17(3):32–39, 2015.
- [279] PETRAS Internet of Things Research Hub. Developing a consumer security index for consumer IoT devices (CSI). <https://www.petrashub.org/portfolio-item/developing-a-consumer-security-index-for-domestic-iot-devices-csi/>, September 2018.
- [280] Ian Phau, Marishka Sequeira, and Steve Dix. Consumers’ willingness to knowingly purchase counterfeit products. *Direct Marketing: An International Journal*, 2009.
- [281] Todd Powers, Dorothy Advincula, Manila S Austin, Stacy Graiko, and Jasper Snyder. Digital and social media in the purchase decision process: A special report from the advertising research foundation. *Journal of advertising research*, 52(4):479–489, 2012.
- [282] Lee Rainie, Sara Kiesler, Ruogu Kang, and Mary Madden. Anonymity, Privacy, and Security Online, September 2013.
- [283] Paul Resnick, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman. Reputation systems. *Commun. ACM*, 43(12):45–48, December 2000.
- [284] Cate Riegner. Word of mouth on the web: The impact of web 2.0 on consumer purchase decisions. *Journal of advertising research*, 47(4):436–447, 2007.
- [285] Karen Rose, Scott Eldridge, and Lyman Chapin. The internet of things: An overview. *The Internet Society (ISOC)*, 80, 2015.
- [286] Joel Ross, Andrew Zaldivar, Lilly Irani, and Bill Tomlinson. Who are the turkers? worker demographics in amazon mechanical turk. *Department of Informatics, University of California, Irvine, USA, Tech. Rep*, 2009.
- [287] Beate Rössler. *The value of privacy*. Polity, Cambridge, UK ; Malden, MA, english ed edition, 2005.
- [288] Beate Rossler. *The value of privacy*. John Wiley & Sons, 2018.
- [289] Norman Sadeh, Jason Hong, Lorrie Cranor, Ian Fette, Patrick Kelley, Madhu Prabaker, and Jinghai Rao. Understanding and capturing people’s privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing*, 13(6):401–412, 2009.
- [290] Naveed Saif, Nasir Razzaq, Muhammad Amad, and Sajid Gul. Factors affecting consumers’ choice of mobile phone selection in pakistan. *European Journal of Business and Management*, 4(12):16–26, 2012.
- [291] Johnny Saldaña. *The coding manual for qualitative researchers*. Sage, 2015.
- [292] Mesay Sata. Factors affecting consumer buying behavior of mobile phone devices. *Mediterranean Journal of Social Sciences*, 4(12):103, 2013.
- [293] Florian Schaub, Rebecca Balebako, Adam L Durity, and Lorrie Faith Cranor. A design space for effective privacy notices. In *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pages 1–17, 2015.
- [294] Bruce Schneier. *Beyond fear: Thinking sensibly about security in an uncertain world*. Springer Science & Business Media, 2006.

- [295] Holger Schütz and Peter M Wiedemann. Judgments of personal and environmental risks of consumer products—do they differ? *Risk analysis*, 18(1):119–129, 1998.
- [296] Jaydip Sen. Security and privacy issues in cloud computing. In *Cloud Technology: Concepts, Methodologies, Tools, and Applications*, pages 1585–1630. IGI Global, 2015.
- [297] Steve Sheng, Ponnurangam Kumaraguru, Alessandro Acquisti, Lorrie Cranor, and Jason Hong. Improving phishing countermeasures: An analysis of expert interviews. In *2009 eCrime Researchers Summit*, pages 1–15. IEEE, 2009.
- [298] Muzafer Sherif. A study of some social factors in perception. *Archives of Psychology (Columbia University)*, 1935.
- [299] Sabrina Sicari, Alessandra Rizzardi, Luigi Alfredo Grieco, and Alberto Coen-Portisini. Security, privacy and trust in internet of things: The road ahead. *Computer networks*, 76:146–164, 2015.
- [300] Michael Siegrist, George Cvetkovich, and Claudia Roth. Salient value similarity, social trust, and risk/benefit perception. *Risk analysis*, 20(3):353–362, 2000.
- [301] Michael Siegrist, Timothy C Earle, and Heinz Gutscher. Test of a trust and confidence model in the applied context of electromagnetic field (emf) risks. *Risk Analysis: An International Journal*, 23(4):705–716, 2003.
- [302] Vijay Sivaraman, Hassan Habibi Gharakheili, Arun Vishwanath, Roksana Boreli, and Olivier Mehani. Network-level security and privacy control for smart-home IoT devices. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2015 IEEE 11th International Conference on*, pages 163–167. IEEE, 2015.
- [303] Lennart Sjöberg. *Perceived risk vs. demand for risk reduction*. Center for Risk Research, Stockholm School of Economics, 1994.
- [304] Lennart Sjöberg. Factors in risk perception. *Risk analysis*, 20(1):1–12, 2000.
- [305] Lennart Sjöberg and Anders Biel. Mood and belief-value correlation. *Acta Psychologica*, 53(3):253–270, 1983.
- [306] Lennart Sjöberg, Bjørg-Elin Moen, and Torbjørn Rundmo. Explaining risk perception. an evaluation of the psychometric paradigm in risk perception research. *Rotunde publikasjoner Rotunde*, 84:55–76, 2004.
- [307] Michael Warren Skirpan, Tom Yeh, and Casey Fiesler. What’s at stake: Characterizing risk perceptions of emerging technologies. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2018.
- [308] P. Slovic, S. Lichtenstein, and B. Bischhoff. Images of disaster: perception and acceptance of risks from nuclear power. *Electr. Perspect.; (United States)*, 3, 1 1979.
- [309] Jack B Soll and Albert E Mannes. Judgmental aggregation strategies depend on whether the self is involved. *International Journal of Forecasting*, 27(1):81–102, 2011.
- [310] FD Sowby. Radiation and other risks. *Health Physics*, 11(9):879–887, 1965.

- [311] Andrew Speirs-Bridge, Fiona Fidler, Marissa McBride, Louisa Flander, Geoff Cumming, and Mark Burgman. Reducing overconfidence in the interval judgments of experts. *Risk Analysis: An International Journal*, 30(3):512–523, 2010.
- [312] Erin L Spottswood and Jeffrey T Hancock. Should i share that? prompting social norms that influence privacy behaviors on a social networking site. *Journal of Computer-Mediated Communication*, 22(2):55–70, 2017.
- [313] Chauncey Starr. Social benefit versus technological risk. *Science*, pages 1232–1238, 1969.
- [314] Harlan J Strauss and L Harmon Zeigler. The delphi technique and its uses in social science research. *The Journal of Creative Behavior*, 9(4):253–259, 1975.
- [315] James Surowiecki. *The wisdom of crowds*. Anchor, 2005.
- [316] Symantec. Why we need a security and privacy “nutrition label” for IoT devices. <https://www.symantec.com/blogs/expert-perspectives/why-we-need-security-and-privacy-nutrition-label-iot-devices>, February 2019.
- [317] Lisa R Szykman, Paul N Bloom, and Alan S Levy. A proposed model of the use of package claims and nutrition labels. *Journal of Public Policy & Marketing*, 16(2):228–241, 1997.
- [318] L Tanczer, J Blythe, F Yahya, I Brass, M Elsdén, J Blackstock, and M Carr. Summary literature review of industry recommendations and international developments on IoT security. *PETRAS IoT Hub, Department for Digital, Culture, Media & Sport (DCMS)*, 2018.
- [319] Leonie Tanczer, Isabel Lopez Neira, Simon Parkin, Trupti Patel, and George Danezis. Gender and IoT research report. <https://www.ucl.ac.uk/steapp/sites/steapp/files/giot-report.pdf>, November 2018.
- [320] The Center for Information Policy Leadership. Ten steps to develop a multilayered privacy notice. https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/ten_steps_to_develop_a_multilayered_privacy_notice__white_paper_march_2007_.pdf, March 2007.
- [321] The Cyber Security Agency of Singapore. Cybersecurity labelling scheme. <https://www.csa.gov.sg/programmes/cybersecurity-labelling>.
- [322] The Digital Standard. The standard. <https://www.thedigitalstandard.org/the-standard>.
- [323] Ulf Toelch, Marjolijn J van Delft, Matthew J Bruce, Rogier Donders, Marius TH Meeus, and Simon M Reader. Decreased environmental variability induces a bias for social information use in humans. *Evolution and Human Behavior*, 30(1):32–40, 2009.
- [324] Ilaria Torre, Odnan Ref Sanchez, Frosina Koceva, and Giovanni Adorni. Supporting users to take informed decisions on privacy settings of personal devices. *Personal and Ubiquitous Computing*, 22(2):345–364, 2018.
- [325] Janice Y Tsai, Serge Egelman, Lorrie Cranor, and Alessandro Acquisti. The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research*, 22(2):254–268, 2011.

- [326] Janice Y Tsai, Patrick Kelley, Paul Drielsma, Lorrie Faith Cranor, Jason Hong, and Norman Sadeh. Who's viewed you?: the impact of feedback in a mobile location-sharing application. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2003–2012. ACM, 2009.
- [327] Monique Mitchell Turner, Christine Skubisz, and Rajiv N Rimal. Theory and practice in risk communication: A review of the literature and visions for the future. In *The Routledge handbook of health communication*, pages 174–192. Routledge, 2011.
- [328] Joseph Turow, Lauren Feldman, and Kimberly Meltzer. Open to exploitation: America's shoppers online and offline. *Departmental Papers (ASC)*, page 35, 2005.
- [329] Joseph Turow, Jennifer King, Chris J Hoofnagle, Amy Bleakley, and Michael Hennessy. Contrary to what marketers say, americans reject tailored advertising and three activities that enable it. *Retrieved October, 14:2009*, 2009.
- [330] Matt Twyman, Nigel Harvey, and Clare Harries. Trust in motives, trust in competence: Separate factors determining the effectiveness of risk communication. *Judgment and Decision Making*, 3(1):111, 2008.
- [331] UL. Identity management & security. <https://ims.ul.com/IoT-security-rating>.
- [332] UL. Methodology for marketing claim verification: Security capabilities verified to level bronze/silver/gold/platinum/diamond, UL MCV 1376. <https://www.shopulstandards.com/ProductDetail.aspx?UniqueKey=35953>.
- [333] Underwriters Laboratories. IoT security rating. empowering consumers through product security labeling. <https://ims.ul.com/IoT-security-rating>.
- [334] Underwriters Laboratories. Verify. <https://verify.ul.com/search>.
- [335] European Union. Energy efficient products. <https://ec.europa.eu/energy/en/topics/energy-efficiency/energy-efficient-products>, 2017.
- [336] United States Census Bureau. Educational attainment in the United States: 2018. <https://www.census.gov/data/tables/2018/demo/education-attainment/cps-detailed-tables.html>, 2018.
- [337] Blase Ur, Pedro Giovanni Leon, Lorrie Faith Cranor, Richard Shay, and Yang Wang. Smart, useful, scary, creepy: perceptions of online behavioral advertising. In *proceedings of the eighth symposium on usable privacy and security*, page 4. ACM, 2012.
- [338] U.S. Senate Committee on Commerce, Science, & Transportation. Strengthening the cybersecurity of the internet of things. <https://www.commerce.senate.gov/public/index.cfm/hearings?ID=A6113AB7-E89B-48C7-B555-E3CBB1466040>, April 2019.
- [339] Will Usher and Neil Strachan. An expert elicitation of climate, energy and economic uncertainties. *Energy policy*, 61:811–821, 2013.
- [340] Trisha Van Zandt and Roger Ratcliff. Statistical mimicking of reaction time data: Single-process models, parameter variability, and mixtures. *Psychonomic Bulletin & Review*, 2(1):20–54, 1995.

- [341] Duco Veen, Diederick Stoel, Mariëlle Zondervan-Zwijenburg, and Rens Van de Schoot. Proposal for a five-step method to elicit expert judgment. *Frontiers in psychology*, 8:2110, 2017.
- [342] Detlof Von Winterfeldt, Richard S John, and Katrin Borchering. Cognitive components of risk ratings. *Risk Analysis*, 1(4):277–287, 1981.
- [343] W3C. The platform for privacy preferences 1.0 (p3p1.0) specification. <https://www.w3.org/TR/P3P/#PURPOSE>, April 2002.
- [344] Ari Ezra Waldman. Cognitive biases, dark patterns, and the ‘privacy paradox’. *Dark Patterns, and the ‘Privacy Paradox’ (September 18, 2019)*, 31, 2019.
- [345] Kristen L Walker. Surrendering information through the looking glass: Transparency, trust, and protection. *Journal of Public Policy & Marketing*, 35(1):144–158, 2016.
- [346] Guijing Wang, Stanley M Fletcher Carley, and H Dale. Consumer factors influencing the use of nutrition information sources. *ACR North American Advances*, 1995.
- [347] Frederick E Webster. Is industrial marketing coming of age? *Review of marketing*, pages 138–59, 1978.
- [348] Zack Whittaker. Many smart home device makers still won’t say if they give your data to the government. <https://techcrunch.com/2019/12/11/smart-home-tech-user-data-government/>, 2019.
- [349] Alexander Wieneke, Christiane Lehrer, Raphael Zeder, and Reinhard Jung. Privacy-related decision-making in the context of wearable use. In *PACIS*, page 67, 2016.
- [350] Charlie Wilson, Tom Hargreaves, and Richard Hauxwell-Baldwin. Benefits and risks of smart home technologies. *Energy Policy*, 103:72–83, 2017.
- [351] Angela G Winegar and Cass R Sunstein. How much is data privacy worth? a preliminary investigation. *Journal of Consumer Policy*, 42(3):425–440, 2019.
- [352] Ashleigh Wood. Privacy notices: Make yours the best in show. <https://www.smartinsights.com/marketplace-analysis/digital-marketing-laws/privacy-notices-make-best-show/>, May 2016.
- [353] Daniel Wood, Noah Apthorpe, and Nick Feamster. Cleartext data transmissions in consumer IoT medical devices. In *Proceedings of the 2017 Workshop on Internet of Things Security and Privacy*, pages 7–12. ACM, 2017.
- [354] World Wide Web Consortium. Open interconnect consortium. <https://www.w3.org/TR/wot-thing-description/>.
- [355] World Wide Web Consortium. Open interconnect consortium. <https://www.w3.org/TR/wot-security/#wot-privacy>.
- [356] Brian Wynne. Institutional mythologies and dual societies in the management of risk. In *The risk analysis controversy*, pages 127–143. Springer, 1982.
- [357] Ilan Yaniv and Eli Kleinberger. Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational behavior and human decision processes*, 83(2):260–281, 2000.

- [358] YourThings. Evaluating and scoring smart-home devices to improve security! <https://yourthings.info/about>.
- [359] Tianlong Yu, Vyas Sekar, Srinivasan Seshan, Yuvraj Agarwal, and Chenren Xu. Handling a trillion (unfixable) flaws on a billion devices: Rethinking network security for the internet-of-things. In *Proceedings of the 14th ACM Workshop on Hot Topics in Networks*, page 5. ACM, 2015.
- [360] Valarie A Zeithaml, Leonard L Berry, and Ananthanarayanan Parasuraman. The behavioral consequences of service quality. *Journal of marketing*, 60(2):31–46, 1996.
- [361] Eric Zeng, Shrirang Mare, and Franziska Roesner. End user security & privacy concerns with smart homes. In *Symposium on Usable Privacy and Security (SOUPS)*, 2017.
- [362] Serena Zheng, Marshini Chetty, and Nick Feamster. User perceptions of privacy in smart homes. *arXiv preprint arXiv:1802.08182*, 2018.
- [363] Tao Zhou. An empirical examination of initial trust in mobile banking. *Internet Research*, 21(5):527–540, 2011.
- [364] Feng Zhu and Xiaoquan Zhang. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of marketing*, 74(2):133–148, 2010.
- [365] Jan Henrik Ziegeldorf, Oscar Garcia Morchon, and Klaus Wehrle. Privacy in the internet of things: threats and challenges. *Security and Communication Networks*, 7(12):2728–2742, 2014.

Appendix A

Survey from “Privacy Expectation and Preferences...”

A.1 Sample Survey Scenario

You are at a **friend’s house**. All rooms have **presence sensors that are used to determine when to switch on and off the lights to reduce costs and save energy**. You are **not told how long the data will be kept**.

1. This use of my data would be beneficial to me.
 - Strongly disagree
 - Disagree
 - Neither agree nor disagree
 - Agree
 - Strongly agree
2. I think scenarios like this happen today.
 - Strongly disagree
 - Disagree
 - Neither agree nor disagree
 - Agree
 - Strongly agree
3. I think scenarios like this will happen within 2 years.
 - Strongly disagree
 - Disagree
 - Neither agree nor disagree
 - Agree
 - Strongly agree
4. (If “disagree” or “strongly disagree” for Q3) I think scenarios like this will happen within 10 years.
 - Strongly disagree
 - Disagree

- Neither agree nor disagree
 - Agree
 - Strongly agree
5. How would you feel about the data collection in the situation described above if you were not told with whom the data would be shared, how long it would be kept or how long it would be used for?
- Very comfortable
 - Comfortable
 - Neither comfortable nor uncomfortable
 - Uncomfortable
 - Very uncomfortable
6. How would you feel about the data collection in the situation described above if you were given no additional information about the scenario?
- Very comfortable
 - Comfortable
 - Neither comfortable nor uncomfortable
 - Uncomfortable
 - Very uncomfortable
7. I would want my mobile phone to notify me every time this data collection occurs.
- Strongly disagree
 - Disagree
 - Neither agree nor disagree
 - Agree
 - Strongly agree
8. I would want my mobile phone to notify me only the first time this data collection occurs.
- Strongly disagree
 - Disagree
 - Neither agree nor disagree
 - Agree
 - Strongly agree
9. I would want my mobile phone to notify me every once in a while when this data collection occurs.
- Strongly disagree
 - Disagree
 - Neither agree nor disagree
 - Agree
 - Strongly agree
10. If you had the choice, would you allow or deny this data collection?
- Allow
 - Deny

A.2 Summary Questions

1. Keeping in mind the 14 scenarios, how often would you be interested in seeing a summary of all such data collection?
 - Every day
 - Every month
 - Every year
 - Never
2. Keeping in mind the 14 scenarios, what would make you comfortable with sharing data in such situations? [text entry]
3. Keeping in mind the 14 scenarios, what would make you uncomfortable with sharing data in such situations? [text entry]

A.3 IUIPC Questions

1. Consumer online privacy is really a matter of consumers' right to exercise control and autonomy over decisions about how their information is collected, used, and shared.
 - Strongly agree
 - Somewhat agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Somewhat disagree
 - Strongly disagree
2. Consumer control of personal information lies at the heart of consumer privacy.
3. I believe that online privacy is invaded when control is lost or unwillingly reduced as a result of a marketing transaction.
 - Strongly agree
 - Somewhat agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Somewhat disagree
 - Strongly disagree
4. Companies seeking information online should disclose the way the data are collected, processed, and used.
 - Strongly agree
 - Somewhat agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Somewhat disagree

- Strongly disagree
5. A good consumer online privacy policy should have a clear and conspicuous disclosure. It is very important to me that I am aware and knowledgeable about how my personal information will be used.
- Strongly agree
 Somewhat agree
 Agree
 Neither agree nor disagree
 Disagree
 Somewhat disagree
 Strongly disagree
6. It usually bothers me when online companies ask me for personal information.
- Strongly agree
 Somewhat agree
 Agree
 Neither agree nor disagree
 Disagree
 Somewhat disagree
 Strongly disagree
7. When online companies ask me for personal information, I sometimes think twice before providing it.
- Strongly agree
 Somewhat agree
 Agree
 Neither agree nor disagree
 Disagree
 Somewhat disagree
 Strongly disagree
8. It bothers me to give personal information to so many online companies.
- Strongly agree
 Somewhat agree
 Agree
 Neither agree nor disagree
 Disagree
 Somewhat disagree
 Strongly disagree
9. I'm concerned that online companies are collecting too much personal information about me.
- Strongly agree
 Somewhat agree
 Agree
 Neither agree nor disagree

- Disagree
- Somewhat disagree
- Strongly disagree

A.4 Demographic Questions

1. How old are you? [text entry]
2. What is your gender?
 - Female
 - Male
 - Other
 - Prefer not to answer
3. What is the highest degree you have earned?
 - No high school degree
 - High school degree
 - College degree
 - Professional degree (masters/PhD)
 - Associates degree
 - Medical degree
 - Prefer not to answer
4. What is your income range?
 - Less than \$15,000/ year
 - \$15,000/ year - \$24,999/year
 - \$25,000/ year - \$34,999/ year
 - \$35,000/ year - \$49,999/ year
 - \$50,000/ year - \$74,999/ year
 - \$75,000/ year - \$99,999/ year
 - \$100,000/ year - \$149,999/year
 - \$150,000/year - \$199,999/ year
 - \$200,000/ year and above
 - Prefer not to answer

Appendix B

Survey from “The Influence of Friends and Experts...”

B.1 Survey Scenarios

The following is the list of scenarios that were presented to the participants in the control condition. We had three *allow* scenarios (A1 – A3), three *deny* scenarios (D1 – D3), and three *balanced* scenarios (B1 – B3). The actual order of the scenarios were randomized for each participant.

- (A1) You are at a department store. This message is displayed on your smartphone: This store has temperature sensors that check for abnormal temperatures, which indicate potential hazards, e.g., fire. This data will be kept for one day.
- (A2) You are at work. This message is displayed on your smartphone: This building has temperature sensors that check for abnormal temperatures, which indicate potential hazards, e.g., fire. This data will be kept for one day.
- (A3) You are at a library. This message is displayed on your smartphone: This library has presence sensors in each room that are used to determine when to switch the lights on and off to reduce costs and save energy. This data will be kept until the room is no longer occupied.
- (D1) You are at a department store. This message is displayed on your smartphone: This store has a facial recognition system that takes pictures of customers’ faces automatically as they enter the store in order to identify returning customers. This method is used to keep track of your orders and make suggestions based on your purchasing habits. Your picture will never be deleted.
- (D2) You are at a library. This message is displayed on your smartphone: This library has an iris scanner that scans customers’ irises automatically as they enter the library in order to identify returning visitors. This is used to keep track of your visits and make suggestions based on your habits. Your iris scan will never be deleted.
- (D3) You are in a public restroom. This message is displayed on your smartphone: This restroom has cameras that are recording video of the entire room. The video is shared with law enforcement to improve public safety. This video will never be deleted.
- (B1) You are at the library. This message is displayed on your smartphone: Your smartwatch is

keeping track of your specific position. Your position is used by the smartwatch to determine possible escape routes in the case of an emergency. This data will never be deleted.

- (B2) You are at work. This message is displayed on your smartphone: This building uses fingerprint scanners instead of keys to unlock office doors and the break room door. This data is also used to track where employees are in the building. Your fingerprint data will never be deleted.
- (B3) You are in a public restroom. This message is displayed on your smartphone: This restroom has presence sensors to detect whether someone is present. This data is shared with law enforcement to improve public safety and they will keep it for one year.

B.2 Sample Survey Questions

These are the questions that we asked the participants in the experimental condition, which included consistent social cues from privacy experts in the scenarios. Here is a sample scenario in this condition:

- (A1) You are at a department store. This message is displayed on your smartphone: This store has temperature sensors that check for abnormal temperatures, which indicate potential hazards, e.g., fire. This data will be kept for one day. More than 85% of privacy experts allowed this data collection.

B.2.1 Questions Posed at the End of Each Scenario

1. What type of data is being collected in the scenario? (In three scenarios, we asked about the data type with the following choices)
 - Video
 - Audio
 - Specific position
 - Presence
 - Temperature
 - Fingerprint
 - Image of iris
 - Image of face
 - Other (please specify) [text entry].(In three other scenarios, we asked about the location of data collection with the following choices)
 - Coffee shop
 - Workplace
 - Home
 - Library
 - Public restroom
 - School
 - Department store

- Other (please specify) [text entry]
- (In the remaining three scenarios, we asked about the retention time and the choices were)
- 1 day
 - 1 week
 - 6 months
 - 1 year
 - Until the room is no longer occupied
 - Until you leave
 - Until the end of the shift
 - It will never be deleted
 - Other (please specify) [text entry]
2. If you had the choice, would you allow or deny this data collection?
- Allow
 - Probably allow
 - Probably deny
 - Deny
3. (If the answer to 2 is “Allow” or “Probably allow”) Why would you allow this data collection? (check as many as apply)
- I am comfortable with the type of data being collected
 - I am comfortable with the purpose of data collection
 - I am comfortable with the length of time for which the data is being kept
 - I am comfortable with the location where the data collection is happening
 - I think the data collection is beneficial to me
 - I think the data collection is beneficial to society
 - I don't think the data collection will reveal my identity
 - I think my collected data will be kept securely
 - I don't see any risk in the data collection
 - I don't have enough information to make an informed decision
 - This is what most privacy experts would do
 - This is what most of my friends would do
 - The benefits to me outweigh the risks
 - I think the data collection is required in this situation
 - Other (please specify) [text entry]
4. (If the answer to 2 is “Deny” or “Probably deny”) Why would you deny this data collection? (check as many as apply)
- I am uncomfortable with the type of data being collected
 - I am uncomfortable with the purpose of data collection
 - I am uncomfortable with the length of time for which the data is being kept
 - I am uncomfortable with the location where the data collection is happening
 - I think the data collection is not beneficial to me
 - I think the data collection is not beneficial to society
 - I think the data collection will reveal my identity
 - I think my collected data will not be kept securely

- I see potential risks in the data collection
 - I don't have enough information to make an informed decision
 - This is what most privacy experts would do
 - This is what most of my friends would do
 - The risks outweigh the benefits to me
 - I don't think the data collection is required in this situation
 - Other (please specify) [text entry]
5. This use of my data would be beneficial to me.
- Strongly agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Strongly disagree
6. This use of my data would be beneficial to society.
- Strongly agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Strongly disagree
7. Regardless of whether you would allow or deny the data collection, how confident are you that this was the right decision for you?
- Extremely confident
 - Moderately confident
 - Somewhat confident
 - Only slightly confident
 - Not at all confident

B.2.2 Questions Posed at the End of Nine Scenarios

1. When considering the 9 scenarios above, how much were you influenced by the decisions that privacy experts made in these scenarios? (we asked about friends in the conditions in which we showed the social cues from friends)
 - Very influenced to do what the experts did
 - Slightly influenced to do what the experts did
 - Not at all influenced
 - Slightly influenced to do opposite of what the experts did
 - Very influenced to do opposite of what the experts did
2. (If the answer to 1 is "Very influenced to do what the experts did" or "Slightly influenced to do what the experts did") What are the reason(s) you were influenced to do what the privacy experts did when deciding to allow or deny the data collection? (check as many as apply)
 - I didn't have a strong opinion about allowing or denying the data collection

- I generally trust privacy experts when making this kind of decision
 - I think privacy experts have more technical knowledge about the data collection
 - I think privacy experts have more background information about the data collection
 - I usually agreed with the actions that were taken by the privacy experts in this survey
 - I generally like to find out what other people have done when making a decision
 - It is easier to do what other people have done than to make the decision on my own
 - I am not sure why I was influenced
 - Other (please specify) [text entry]
3. (If the answer to 1 is “Very influenced to do opposite of what the experts did” or “Slightly influenced to do opposite of what the experts did”) What are the reason(s) you were influenced to do the opposite of what the privacy experts did when deciding to allow or deny the data collection? (check as many as apply)
- I didn’t have a strong opinion about allowing or denying the data collection
 - I generally don’t trust privacy experts when making this kind of decision
 - I think I have more technical knowledge about the data collection
 - I think I have more background information about the data collection
 - I usually disagreed with the actions that were taken by the privacy experts
 - I generally like to do the things that are different from what other people do
 - I am not sure why I was influenced
 - Other (please specify) [text entry]
4. (If the answer to 1 is “Not at all influenced”) What are the reason(s) you were not influenced by privacy experts’ actions when deciding to allow or deny the data collection? (check as many as apply)
- I didn’t have a strong opinion about allowing or denying the data collection
 - I generally don’t trust privacy experts when making this kind of decision
 - I think I have more technical knowledge about the data collection
 - I think I have more background information about the data collection
 - I generally make decisions on my own
 - I make these kinds of decisions on my own
 - I usually disagreed with the actions that were taken by the privacy experts in this survey
 - I would want to know more about the people whose actions are being shown to me before I would trust them
 - I am not sure why I wasn’t influenced
 - Other (please specify) [text entry]

(Only for the last scenario: e.g., if the last scenarios was A1)

This is the last scenario you were shown: You are at a department store. This message is displayed on your smartphone: This store has temperature sensors that check for abnormal temperatures, which indicate potential hazards, e.g., fire. This data will be kept for one day.

1. (We keep the level of consensus as before and change the influencers from privacy experts to friends or from friends to privacy experts) If you were told that more than 85% of your friends who use this app allowed the data collection in this scenario, would you allow or deny this data collection?
- Allow

- Probably allow
 - Probably deny
 - Deny
2. (We keep the influencers the same and change the consensus level to the opposite majority decision from more than 85% to fewer than 15% or from more than 65% to fewer than 35%) If you were given the same scenario but told that fewer than 15% of privacy experts allowed the data collection, would you allow or deny this data collection?
- Allow
 - Probably allow
 - Probably deny
 - Deny
3. We have previously shown you how privacy experts and your friends who use this app acted in similar situations. Who are the other people or organizations whose actions would influence yours in scenarios like these? Which would be most influential? [text entry]
4. For each type of person described below, please specify your level of agreement with the following statement: I would trust [blank] to give me good advice when I need to make a decision about allowing devices to collect and use my information. *Choices for blank are:*
- Privacy experts
 - My family
 - My real-life friends
 - People working in technical fields
 - My colleagues
 - My social network friends
 - No one except myself
- Strongly agree
 - Agree
 - Neither agree nor disagree
 - Disagree
 - Strongly disagree
 - Not applicable
5. Please specify your level of agreement with the following statements.
- I think privacy experts have more technical knowledge about the data collection than I do.
 - I think my friends have more technical knowledge about the data collection than I do.
 - I think privacy experts have more background information about the data collection than I do.
 - I think my friends have more background information about the data collection than I do.
 - I generally like to find out what other people have done when making a decision.

- It is easier to do what other people have done than to make the decision on my own.
 - I generally make decisions on my own.
 - I would want to know more about the people whose actions are being shown to me before I would trust them.
- Strongly agree
 Agree
 Neither agree nor disagree
 Disagree
 Strongly disagree
 Not applicable
6. Please specify your level of agreement with the following statements.
- I have sufficient knowledge about privacy to make a decision about allowing my information to be collected and used.
 - I have sufficient knowledge about the technologies mentioned in the scenarios to make a decision about allowing my information to be collected and used.
 - The scenarios generally provided sufficient information about how data would be used to make a decision about allowing my information to be collected and used.
- Strongly agree
 Agree
 Neither agree nor disagree
 Disagree
 Strongly disagree
 Not applicable
7. What qualities would make you likely to be influenced by a specific group of people when you need to make decisions like the ones in our scenarios? (check as many as apply)
- Having some background in technology
 Being related to them by blood, i.e. family members
 Having some friendship history with them
 Knowing them well
 Being close friends or family
 Being reliable
 Being honest
 Caring about me
 Having no ulterior motive
 Other (please specify)

B.2.3 Demographic Questions

1. What is your age? [text entry]
2. What is your gender?
 Male

- Female
 - Other
 - Prefer not to answer
3. What is the highest degree you have earned?
- No high school degree
 - High school degree
 - College degree
 - Professional degree (masters/PhD/medical/law)
 - Associates degree
 - Prefer not to answer
4. What is your income range?
- Less than \$25,000/year
 - \$25,000/year - \$49,999/year
 - \$50,000/year - \$74,999/year
 - \$75,000/year - \$99,999/year
 - \$100,000/year - \$124,999/year
 - \$125,000/year - \$149,999/year
 - \$150,000/year - \$174,999/year
 - \$175,000/year - \$199,999/year
 - \$200,000/year and above
 - Prefer not to answer

Appendix C

Interview Script and Codebook from “Exploring How Privacy and Security...”

C.1 Screening Survey Questions

1. What is your full name? [text entry]
2. What is your email address to contact you? [text entry]
3. What IoT device(s) do you have? [text entry]
4. How did you get your IoT device(s)? (check as many as apply)
 - I purchased it/them online
 - I purchase it/them in store
 - Somebody gave it/them to me as gift
 - Somebody in my house purchased it/them
 - Other (please specify) [text entry]
5. What time of the day are you available for the interview? [text entry]
6. What is your age? [text entry]
7. What is your gender?
 - Female
 - Male
 - Other
8. What is the highest degree you have earned?
 - No high school degree
 - High school degree
 - College degree
 - Professional degree (masters/PhD/medical/law)
 - Associates degree
 - Prefer not to answer
9. What is your current employment status?
 - Full-time employment

- Part-time employment
 - Unemployed
 - Self-employed
 - Home-maker
 - Student
 - Retired
 - Prefer not to answer
10. What is your occupation? [text entry]
 11. Do you have a degree in computer science or related fields (if you have, please specify what that degree is)? [text entry]
 12. Do you have a technical background? (if yes, please specify what your background is)? [text entry]

C.2 Interview Questions

C.2.1 Questions about Electronic Devices

1. Can you please list the last two electronic devices you bought?
2. When was the last time you compared multiple brands of the same product when making a purchase? What was the product? What were the factors that you considered when making the purchasing decision to buy or not to buy the product?
3. How frequently do you make comparisons like this when purchasing?
4. Are these sorts of factors the ones that you always consider or are there any other factors that you may consider for other types of products before making a purchasing decision?
5. When you are purchasing electronic devices such as camera, thermostat, smartphone, toaster or TV when do you like to buy them in-store and when do you like to buy them online?

C.2.2 Questions about IoT Devices

1. How do you define Internet of Things (IoT)?
2. What are some main requirements that you believe a device should have to be considered as an IoT device?
3. What does the phrase smart home mean to you?
4. What are some IoT devices associated with a smart home?
5. You've previously specified that you have purchased at least one IoT device, what were they? How did you buy them? How long have you owned each of these devices?
6. What made you buy your smart device(s)?
7. Have you ever considered buying a smart device, but you ended up not buying it? What was the device? What made you not buy the device?

8. What was the last IoT device that you bought? Did you buy it online or in store and why?
9. (if they have more than one IoT device) Is this the same process that you use for all of your IoT devices or have you used other processes? How did you make this decision to buy them online/store?
10. What are the factors that you considered when purchasing your last IoT device? Did you compare different devices based on these factors? How often do you do comparison shopping regarding IoT devices?
11. (if they have more than one IoT device) Are these factors specific for this device or are they the same factors that you considered when purchasing your other IoT devices?
12. Which of these factors do you consider to be more important to you? Why?
13. Have you ever experienced any issue or have you had any concern toward your IoT device(s)? What were those concerns? How did you manage them?
14. How do you define privacy regarding IoT devices?
15. Have you ever had any privacy-related concern about your IoT device(S)? What were you concerned about? How did you manage that concern?
16. How do you define security regarding IoT devices?
17. Have you ever had any security-related concern about your IoT device(S)? What were you concerned about? How did you manage that concern?
18. How important is it for you to know the privacy and security information of your smart device while making the purchasing decision? Why?
19. Everything else being equal, would you pay more for a device that had privacy and security information provided as compared to one that did not? Why?
20. How comfortable are you with the data collection of your IoT device(s)?
21. Have you ever read the privacy policy of your device(s)? Why or why not?
22. How much do you think you know about the privacy-related information of your device(s)?
23. What do you want to know most about the privacy and security of your IoT devices?
24. How would you rank these items based on their impact on your IoT-related purchasing decisions? (cards are privacy information, security information, brand, price, and all the other factors participant mentioned throughout the interview)

C.2.3 Questions about Label Evaluation

1. Imagine you want to buy a smart thermostat, please take a look at these three labels. They are all the same price and have the same features. Tell me which one would you buy? also tell me how did you make this decision or which piece or pieces of information helped you to make the decision?
2. Read from the first line and tell me what does it convey? Also please circle the information which is vague or not can be presented better.
3. Overall in this label, knowing about which information is more important to you to make

the purchasing decision?

4. In the privacy section, knowing about which information is more important to you and knowing about which information is not very important?
5. Is there any privacy-related information that is missed from this section which you want to know about before making the purchasing decision?
6. In the security section, knowing about which information is more important to you and knowing about which information is not very important?
7. Is there any security-related information that is missed from this section which you want to know about before making the purchasing decision?
8. Is there any other information you think is missing from the label that would help you make a more informed purchasing decision about the device?
9. Have you ever heard of Consumer Reports (CR) or used it by looking at the products' ratings by CR before making a purchase?
10. How influential did you find the CR score and the ratings?
11. What extra information about the ratings do you think can make them more influential?
12. While making a purchasing decision, how useful do you think it is to be presented with privacy and security information of the IoT device(s) in the form of labels? Why would you think this way? In addition to the label, can you think of other useful format to present privacy and security information of IoT devices?
13. If you are purchasing an IoT device in-store, how would you prefer to be presented with this information? Do you want it to be a label on the side of the device's package or do you want it to be available by scanning a QR code on the package? Or do you have other preferences regarding this?
14. If you are purchasing an IoT device online, how would you prefer to be presented with this information? Do you want it to be a label or do you prefer to instead be able to click on a URL that directs you to the privacy policy of the device?
15. How important is it for you to know about publicly reported security vulnerabilities of the device(s) you are interested to buy? How do you think this information would influence your purchasing decisions? What would you wanna know about the vulnerabilities?

C.3 Supplementary Survey Questions

1. Imagine you are deciding between two or more [smart security cameras/smart thermostats/smart toothbrushes] to purchase. How much influence do you think each of the following factors would have on your purchase decision? *Factors are:*
 - Look and feel
 - Customer service
 - Prior experience with the device or similar devices
 - Ease of use

- Reliability
 - Opinion from experts (magazine reviews, electronics store employee)
 - Compatibility with other devices
 - Durability
 - Opinion from friends
 - Opinion from family members
 - Brand
 - Privacy information
 - Security information
 - Customer reviews
 - Price
 - Features
- 1: No influence at all
- 2
- 3
- 4
- 5: A lot of influence
2. Have you ever owned a [smart security camera][smart security cameras/smart thermostats/smart toothbrushes]?
- Yes
- No
3. (if the answer to 2 is “Yes”) How did you get your [smart security camera(s)/smart thermostat(s)/smart toothbrush(s)]? (check as many as apply)
- I purchased it/them online
- I purchase it/them in store
- Somebody gave it/them to me as gift
- Somebody in my house purchased it/them
- Other (please specify) [text entry]
4. Have you ever purchased an IoT device yourself (smart home device or wearable)?
- Yes
- No
5. (if the answer to 4 is “Yes”) What IoT device(s) have you purchased yourself? [text entry]
6. What is your age? [text entry]
7. What is your gender?
- Male
- Female
- Other
8. What is the highest degree you have earned?

- No high school degree
 - High school degree
 - College degree
 - Professional degree (masters/PhD/medical/law)
 - Associates degree
 - Prefer not to answer
9. What is your current employment status?
- Full-time employment
 - Part-time employment
 - Unemployed
 - Self-employed
 - Home-maker
 - Student
 - Retired
 - Prefer not to answer
10. What is your occupation? [text entry]
11. Do you have a degree in computer science or related fields (if you have, please specify what that degree is)? [text entry]
12. Do you have a technical background (if yes, please specify what that degree is)? [text entry]

C.4 Codebook

The following table shows the 8 structural codes and 61 subcodes we used to tag the interview data. The example in front of each code may be tagged with more than one code.

Code	Example
Definition of Smart Device	
Internet connectivity	“Something that can connect to the internet.”
Wifi	“It should be WiFi enabled”
Data enabled	“I guess it could be data enabled”
Bluetooth enabled	“Any device that can communicate wirelessly, through Bluetooth or Wi-Fi”
RFID	“Well, it has to have some way to relate to an app, whether it’s through RFID or Bluetooth, or Internet connection”
Remote controllable	“So you can access it remotely, it has some sort of remote capability.”
Programmable	“Have some rudimentary programmable features”
Interaction with other devices	“So it can be several devices connected together, working together.”
Lifestyle improvement	“it should reduce my, whatever I’m doing, it should make me more productive and efficient.”
Electronic device	“It should be an electronic device, that’s obvious.”
Ability to sense	“it must be able to do some sort of sensing, or meet some sort of parameters that can vary from say, measuring the temperature or . . .”
Physically small	“Also that it needs to be very . . . Physically, it needs to be very small.”
Always on	“Maybe they are functional at all times.”
Easy to use	“Anyone can understand it, kind of fool-proof”
Being smart	“Intelligent electronic devices.”

Code	Example
Personalized	“[having] a profile or account specific to whoever is using it.”
Autonomous	“house that will, automatically adjust its temperature because it will sense what you want”
Learn behaviors	“when you have a really integrated system that you can control and it also learns from your behavior.”
Definition of Privacy	
Collected data	“I think privacy, I need to know what information it’s collecting ”
Data being shared	“whether or not it will send it somewhere else”
Purpose of data collection	“So, the idea that if it’s collecting my data, the data, or my data they tell me what they do with the data”
Identifiability	“So if you can identify me by name or location or some other significant identifier, then I have privacy concerns.”
Retention time	“how long do they keep it? ’Cause I can only view it for so many days, but then do they keep it?”
Inferred data	“ I think that obviously, there are things that might be inferred from usage, but there’s sort of a trade-off where some of the things that make your device smart require those inferences be made”
Choice and control	“privacy it’s whether it’s up to me or them how they use my data.”
Definition of Security	
Hack or unauthorized access	“security is about how they like, for example what if somebody hacks into the system and do that”
Passwords and authentication	“ I wanted to make sure that I can change a password on it, so that it’s not a default”
Firewalls	“we have buffers and firewalls and all that sort of thing. That’s security.”
Lock	“Security is, we have our servers in a super locked-down location with no physical access”
Encryption	“Encryption, I guess, is the only word that I would think of”
Risks associated with unauthorized use of data	“security is what happens to that information after they have it and my risk for something.”
Reasons to Buy Smart Home Devices	
Safety	“this particular thing is for safety reasons.”
Saving energy	“So I want to be able to save energy by being able to program the lights to turn themselves off.”
Convenience	“the rest is ... the lighting and stuff is just convenience.”
Curiosity to experience the technology	“Curiosity, honestly.”
Price	“Well, we were with my mother-in-law and there was like a two-for-one deal, so that was probably why we got it right that second”
Lifestyle improvement	“I have a couple kids, my wife works from home, so our schedule wasn’t traditional in the sense that I could just turn it off at 7:00 when we left for the day and turn it back on or set it to come back on, so I wanted one that would learn our schedule. Not only during the week, my wife works four 10s, so on her off day, she’s usually out and about so we don’t need it throughout, so it was a more complex schedule than what I wanted to sit there and program into a thermostat.
Fun	“That could be fun, but it’s not a necessity”
Necessity	“That it wasn’t something I felt was gonna replace something else. And it wasn’t gonna be, take over my life or anything, but ... So now a necessity. Now we have it and I do ... Now we have a house, we have two floors, and so sometimes I’m upstairs and I wanna ask the Google Home something and it’s downstairs.”
Reasons to Buy Wearables	

Code	Example
Lifestyle/health improvement	“I was trying to improve sleep hygiene by being able to leave my phone outside the bedroom that night. I didn’t want to have to set an alarm by looking at a screen.”
Curiosity to experience the technology	“More just curiosity. I mean, I do want to get it in shape and I thought it would be partially a reminder of that, in a way, because I would notice it when it was on my wrist. But, mostly I guess just the novelty of it, just the curiosity.”
Convenience	“You use your Apple watch or your Samsung watches, depending on how your lifestyle is. Your music, your choices. It’s ease and convenience.”
Price	“I only bought it because it was cheap. Not cheap, but half price.”
Necessity	“For a long time, I think it was a necessity and then I just stopped using it for a little bit because I felt too tied to it. Now, it’s clearly not a necessity, but it’s definitely something that when I don’t have it, I miss it.”
Reasons Not to Buy Smart Home Devices	
Lack of necessity	“just the fact that I didn’t really think that I needed it”
Price	“The main reason why I didn’t buy it was ... there are not terribly expensive but it was expensive enough that I didn’t buy it just because it wasn’t really worth it to me, so cost was kind of a factor.”
Security concern	“I wasn’t confident in the security of it”
Privacy concern	“I don’t like Alexa because it’ll sometimes come on when I don’t have any request for it. So, it’s always listening and I don’t like that.”
Device/electronic failures	“I was looking at a new garage door opener. I was very close to buying a smart garage door opener, but there was only a couple brands and they didn’t have the best reviews at this point, like they were still glitchy and didn’t work as well as people would expect them to, so I decided not to.”
Not improving the lifestyle/less convenience	“I found it was less convenient to get my phone and open the app and turn the light off, than just go do it.”
Reasons Not to Buy Wearables	
Lack of necessity	“I didn’t have a reason to buy it yet.”
Price	“Because, you would’ve also had to have the iPhone and they’re like \$1,000.00, and it just wasn’t worth it for the novelty of it.”
Features	“it doesn’t seem like it would do that much for me.”
Device failures	“Yeah, the main options I was looking at were the Fitbits and at the time there was a lot of discussion online about them failing. Then about a year in have a dead device. So it just didn’t seem like the technology was there yet.”
Aesthetic reasons	“I can’t decide. Actually, they’re too big. I don’t like a bulky watch”
Online Shopping Versus In-Store Shopping	
Convenience	“Yeah, online. Plus ... I would say it’s all convenience. I don’t feel like I should have to go to a store anymore.”
Price/deal	“Well, if the price is cheaper online, I would get it online, and often it is.”
Look and feel	“Televisions and stuff, I like to go see them. I like to see televisions and refrigerators and stuff like that”
Urgency	“but if it was in a store because I need them immediately.”
Opinion from expert	“Well, Apple I had been looking online. Like, I had already done some research online and then I went in and talked to, like, a genius bar and staff.”
Ability to compare/more variety	“And often, in stores, any more for things that aren’t ... Their main things that they sell have really limited selection and it’s usually not what I want.”

Table C.1: Codebook we used for our interview study.

Appendix D

Interview Scripts, Surveys, and Codebook from “Ask the Experts: What...”

Factors	Reason(s) to include	Reason(s) not to include
Almost all of the experts wanted to include		
Type of data that is being collected	Key factor that affects consumer purchase decisions.	R1: Challenging to determine what data types should be included on the label as there is a trade-off between making them understandable using generalization or making them highly-technical with specific details.
Who the data is shared with	R1: Encourage companies to share data with fewer entities. R2: Informs consumers about how their data will be used.	This information is dynamic. Therefore, it’s hard for the companies to specify all the parties upfront.
Who the data is sold to	R1: It’s important for consumers to have information about data-driven business model of companies. R2: Informs consumers about how their data will be used.	R1: Consumers won’t be able to recognize the name of the parties their data is sold to. R2: Companies may not be able to anticipate all the parties that data will be sold to.
Retention time	R1: Encourages companies to keep the data for a shorter amount of time. R2: Helps consumers make more informed privacy-related decisions.	Since consumers can’t force the companies to retain the data as promised, this information is not useful for them.
Whether or not the collected data will be linked with data obtained from other sources	R1: Currently it’s not clear to consumers whether or not the collected data is being linked with internal or external sources. R2: Helps consumers make an informed purchase decision. R3: Indicator of the data-driven business model of the company.	Does not convey risk to consumers.
Whether or not the device is getting cryptographically signed and critical automatic security updates	R1: Consumers understand the importance of automatic security updates. R2: New vulnerabilities in software are discovered very often and hence continuous checking for updates is needed. R3: In both manufacturers and consumers best interest. To manufacturer as a sign of support and longevity and to consumers as a sign of the lifetime of the device.	R1: Does not guarantee a well implemented and trustworthy update process. R2: Sometimes there are reasons not to have automatic security updates. These reasons include legal considerations or if the device is incapable of receiving updates. R3: Automatic updates should be a requirement and there is no need to explicitly mention them.
Security rating for the device from an independent security assessment organization	Rating provides a succinct, understandable, and reliable way to learn about the security of the device, from experts’ point of view.	R1: The security level of a device can change frequently. Therefore, the rating will also need to change as new vulnerabilities are discovered. R2: Puts a pressure on the companies to focus on getting full-star rating instead of improving their security practices. R3: Currently the metric for the rating does not exist.
Type of the sensor(s) on the device	Indicates what types of data can be collected and exposed in case of an attack.	Consumers may not understand the capability of some sensors and this information can be complicated to them.
Purpose of data collection	R1: Brings legal accountability to the device manufacturers. R2: Helps consumers make more informed decision about the collection and use of their data.	It’s hard to determine what purposes should be included on the label as there is a trade-off between understandable but useless generalization as well as technical but meaningful details.

Factors	Reason(s) to include	Reason(s) not to include
The date until which security updates will be provided	Critical security information.	For large and well-established companies, long-term commitment would leave them subject to unpredictable expenses. On the other hand, smaller companies (e.g., startups), are likely to make the commitment for updates without much consideration, since they may not be around long enough to have to actually do the updates.
Whether or not the device can still function when data-driven smart features (e.g., the learning function of a smart thermostat) are turned off	It's important for consumers to know whether or not their device can still function without the ability to learn behaviors based on their collected data.	Smartness of IoT devices when the data processing happens on the device does not correlate with privacy.
Whether or not the device has parental control mode	It's important for consumers to know how much control they have.	Little relevance to privacy.
What information can be inferred from the collected data	Critical privacy information	It's hard for the companies to predict and specify what can be inferred from the collected data in short and long term.
Where the collected data is processed (local, cloud, and hub as well as geographic location)	Depending on the countries that the data is processed at, different regulations and laws may be applicable.	R1: The data will be processed in all three locations (cloud, local, and hub). Therefore, this information is not useful for consumers. R2: Knowing where the data is processed at is not an indicator of privacy or security.
Most of the experts wanted to include		
Specifying the access control (e.g., no access control, single-user account, multi-user account) for the device and any accompanying apps	Important to include as the type of the access control is not clear from the type of the device.	R1: This information is mostly a functionality feature of the device, not an indicator of privacy. R2: Consumers may not understand the implications of this factor.
Where the collected data is stored (local, cloud, and hub as well as geographic location)	R1: The legal protection for the data depends on whether the data is stored on the device or on the cloud. R2: Where the data is stored indicates whether or not the data leaves the device. R3: The geographic location of the data may impact legal protections as well as consumer perceptions.	R1: The data will be stored at all three locations (cloud, local, and hub), therefore this information is not useful for consumers. R2: Knowing where the data is stored at is not an indicator of privacy or security.
The control that users are offered (e.g., opt-in, opt-out) for receiving targeted suggestions based on their behavior	Behavioral advertisement leads to many privacy issues. Therefore, it's important to inform consumers about targeted ads and the controls consumer have, if any.	After the data collection has happened, consumers have no choice in controlling what can be learned about their data. Therefore, this information is not useful.
Frequency of data collection (e.g., once a month, on install, every day, hourly, continuous)	Indicates how much detailed information about the consumer is being collected. This factor will be more important if sensitive information is being collected.	Specific frequency of data collection may be needed for the core functionality of the device.
Whether or not the device can still function when Internet connectivity is turned off	Indicates how useful the device is and also whether privacy was being considered as a factor during the design process of the device.	Consumers may not understand the privacy and security implication of this factor.
Listing relevant security and privacy laws and standards to which the manufacturer is complying (e.g., ISO 27001, GDPR)	Increases the accountability for manufacturers.	R1: This can be faked. R2: Current laws are mostly about software and are missing out on hardware. R3: Consumers are not familiar with all the laws and standards.
Listing the standards that the device supports (e.g., Wi-Fi, Zigbee)	R1: Indicates how vulnerable the device is based on the standards it uses. R2: Helps consumers with their purchase decisions.	Does not inform consumers' decision making as consumers may not be familiar with the standards or they may not understand the significance of specific standards.
Privacy rating for the device from an independent privacy assessment organization	R1: This is the most important and understandable information for average consumers to make informed purchase decisions. R2: Adds legitimacy to the label.	Currently the metric for the rating does not exist.

Factors	Reason(s) to include	Reason(s) not to include
Types of physical actuations (e.g., talking, blinking) the device has and in what circumstances they are activated	R1: This information may relate to the physical security of the device. R2: This information indicates what types of data the device can collect. R3: Consumers should know about the types of intrusions that will come with the device before making the purchase.	R1: This information is more relevant to the functionality and does not relate to the privacy of the device. R2: This information does not inform decision making for average consumers. R3: If the data that triggers the actuations is not being kept on the device or is not being transmitted from the device, it will not impose a concern.
Whether or not the device is using any default password	Using a default password is an example of a bad security practice and an indicator of how insecure the device is.	R1: There are some reasons as to why default password is not an indicator of bad security. For example, default passwords could be changeable by consumers. In addition, the importance of the device not using a default password depends on the type of the device. R2: Not having a default password should be a requirement and there is no need to explicitly mention it.
Link to the device's key management (e.g., key storage, key distribution) protocol	Device security can be compromised if keys are compromised.	R1: Consumers as well as experts cannot understand this information. R2: Using specific protocols does not specify the level of security as manufacturers may not know how to configure SSL properly from their device to their server.
Frequency of data sharing (e.g., continuous, on demand, periodic)	R1: The frequency of data sharing is important to be mentioned as this information is not determined by the functionality of the device. R2: The sensitivity of the shared data can be determined by the frequency of data sharing.	R1: The importance of this information depends on the devices and how sensitive the shared data is. R2: The time after which data is released, or how often it is released is not as important as the fact that the data is released at all.
The warranty period of the device	R1: This information informs about the privacy life-cycle. R2: If the warranty covers the security issues, this information could be useful for the physical threats. R3: This information helps consumers in the purchase decisions.	R1: The items that the warranty covers may be only relevant to the functionality, not privacy and security of the device. R2: This information does not inform decision making as the warranty is unrelated to how trustworthy the device is or the ability to use the device after the end of the warranty. R3: If the warranty is voided upon purchase, including the warranty period will reduce users' trust in the devices. Manufacturers may be reluctant to guarantee since the market is volatile.
Granularity of the data being collected, used, and shared (e.g., identifiable, reported in aggregate)	R1: Indicates how easy the individual can be identified through her data. R2: Specifies the risks associated with the collected data.	R1: Currently there is no metric or definition for the de-identification or anonymization of data. R2: This information could be misleading and distract consumers from noticing the consequences of their data when being collected and shared. R3: Companies may view this information as proprietary and be reluctant to disclose it.
Resource usage in terms of power and data (e.g., kw, kbps)	R1: This information can have implications of the safety of the device. R2: This information has environmental implications.	R1: This information does not relate to the safety, privacy, or security of the device. R2: This information does not inform decision making for average consumers.
Special data handling practices for children's data	Data related to children is more sensitive and has specific legal requirements, therefore needs special protection.	Some IoT devices may not be designed for children as their primary users. Therefore, there is no need for them to describe their children's specific data handling practices.
Some of the experts wanted to include		
What authentication methods the device is using	R1: Indicates how vulnerable the device may be to attacks. R2: Indicates how accessible and usable the device is.	R1: Little relevance to privacy or security and could be misleading to consumers. R2: Importance depends on the device.
Sensitivity of the data that is stored/processed (locally, on the cloud, and in the hub)	R1: People will care more about sensitive data, so if they are reading quickly, they can look just at whether there is sensitive data and what is being done with it. R2: There are different legal requirements for the protection of sensitive data (and the relevant laws include definitions of sensitive data for that purpose).	Sensitivity of the information is subjective and there is no metric to measure it.

Factors	Reason(s) to include	Reason(s) not to include
The date of the latest device (firmware) update	This information indicates how vulnerable the device is.	R1: This information does not inform decision making for average consumers as they may not have a correct understanding of the importance of the timing of updates. R2: The latest device update does not convey how secure the device is. In other words, an older software could be more secure than recently updated one, depending on how robust the software was implemented.
Average response time for patching company's products (e.g., within one month)	The average response time indicates the likelihood of them patching new vulnerabilities in the future based on their timely responses to patching known exploits in the past.	R1: The average response times can vary for multiple reasons and this may give the companies an incentive to appear to shorten the average response time without actually improving their security practices. R2: There will be fewer data points for smaller companies to average over.
What protocols the device manufacturer is using to encrypt the data that is being transmitted and stored (on the cloud and locally)	R1: Encryption should follow well-defined standards, and merely stating whether the standard is followed should be enough without more detail. R2: Details of the encryption protocols used would help indicate the level of security protection consumers are being provided with.	R1: The availability of encryption does not guarantee the robustness of the encryption process. For example, encryption may not be implemented correctly or the device may not be using standard protocols. R2: Consumers cannot understand this information.
What network ports the device opens and listens for incoming connections on	R1: Open network ports can show incoming network connections or data being sent out to the Internet. R2: Open network ports indicate the attack surface and this will help to better protect the device.	R1: Network ports are not understandable, practical, or relevant security indicator for consumers.
Whether or not the device is MUD (Manufacturer Usage Description) compliant	MUD indicates the behavior of the device and being MUD compliant is an indicator that the device is acting based on its description.	MUD compliance is volatile and can be violated as soon a new vulnerability is discovered.
Link to the device's known vulnerabilities and when they were reported and patched	R1: Indicates how responsive and trustworthy the company is. R2: This is a critical security information and its importance does not depend on the type of the device.	Companies may be reluctant to disclose the complete list of the device's patched vulnerabilities as that may signal an insecure device.
Where the device was manufactured	R1: Depending on where the device is manufactured in, there may be privacy and security related concerns. R2: Consumers as well as industries may be sensitive to the origin of the device.	R1: Where the device was manufactured is not important. However, it's important to know where the device is being sold. R2: This information does not inform decision making for average consumers. R3: Including this information can help spread the unfair stereotypes.
Link to the manufacturer's anti-tampering practices (e.g., having tamper-detection sensors)	Information about anti-tampering practices indicate how much effort the company has out on securing the device from being physically tampered with.	R1: Physical tampering does not frequently happen to IoT devices. R2: The importance of physical tampering depends on the type of the device. R3: Including information about anti-tampering practices can increase the risk of the device getting attacked.
Periodic security audits	R1: Periodic security audits indicate that the company is paying attention to their security practices. R2: Well known companies complying with periodic security audits will create pressure on even smaller players and the rest of industry to follow suit and comply to audits.	R1: Currently there is no consensus over the definition and metric for security audits. R2: Consumers won't be able to recognize the name of the auditors and which ones are most credible. R3: Knowing whether or not the device is undergoing security audits without mentioning the findings of the audits does not indicate how secure the device.
A few of the experts wanted to include		
List of device-compatible products	R1: Some of the compatible devices can be more concerning than the device itself. R2: This information can be used to signal privacy-protective devices in the ecosystem. R3: It's hard for consumers to keep track of the standards and protocols and this information helps them with their purchase decisions.	R1: This information is more relevant to the functionality and does not relate to the privacy of the device. R2: The list may be too long.
Link to the accompanying app(s) of the device (in case the device has an accompanying app)	This information informs consumers about the availability of the software component of the device, the ratings of the accompanying app(s), and whether or not personal information will be collected.	R1: This information does not relate to the privacy or security of the device. R2: This information does not inform decision making for average consumers.

Factors	Reason(s) to include	Reason(s) not to include
Link to the device's software and hardware bill of material	Providing a link to the software and hardware bill of material enables consumers to check for existing vulnerabilities and if possible, apply appropriate defense mechanisms.	R1: It's hard to list the supply chain and specify the important one to show to consumers R2: Listing software and hardware components does not indicate how secure the device is. R3: Listing the software components can increase the risk of the device getting attacked.
Whether or not the device manufacturer has bug bounty program in place for the device	The availability of the bug bounty program is an indicator of how trustworthy the companies are and much they care about security.	R1: Just knowing the company has a bug bounty program does not convey whether or not the company is taking security seriously and has a process to address the security issues. R2: The cost of having bug bounty program may exceed its benefit and the bug bounty program may attract black-hat hackers.
Where and when the device brand was incorporated	R1: When the brand was incorporated is an important factor as more established brands have more reliable security practices. R2: This information is useful and helps to trace back the vulnerabilities to their makers. R3: It's important to include information about where the device was manufactured in as consumers as well as industries may be sensitive to the origin of the device.	R1: Including this information can help spread the unfair stereotypes. R2: This information does not inform decision making for average consumers.
The score from Consumer Reports	R1: This information helps the trustworthiness of the label by referencing it to a reputed entity like Consumer Reports. R2: This information indicates whether the device and the company are likely to be maintained. R3: This information can increase the security of smart devices across the IoT ecosystem.	R1: This information can be proprietary. R2: This information is subjective. R3: Consumer union does not directly reports on the privacy and security practices of the device.

Table D.1: Most frequent arguments that the experts provided to include or not to include a factor on the label. In this table, R stands for reason.

D.1 Interview Questions with Privacy and Security Experts

You are probably aware that food products have nutrition labels on them that tell you about all the different nutrition facts about them. However, when you buy an IoT device, there is no information like that. We are trying to develop a label for IoT devices that would focus on privacy and security aspects of those devices, similar to nutrition label for foods. Today I am going to talk to you about the things that should be on this privacy and security label for IoT devices.

Before asking about the label, I want you to please:

1. Define security in the context of IoT devices.
2. Can you also define privacy related to the IoT devices.

We are going to talk about what are the different factors that you would like to put on the label. Obviously there is a limited amount of space on the label, but for now don't worry about that and let's just talk about what factors you would like to see.

1. Please specify the security/privacy/general factors that you would want to see on the label.
2. For each factor please specify the levels or values that you think this factor should have.
3. For each factor please specify the choices consumers may be provided with by the manufacturer.
4. Among the factors you mentioned, please specify the ones that are good for experts but probably are not needed for consumers.

(Presenting interviewee with the factors (privacy, security, and general) from the previous experts we have talked to)

5. Please read the factors that other experts provided and compare them with the factors that you mentioned. Let me know if there is any factor, level, or choice that is currently not on your list, but you find it important for consumers to know about and want to add that to your list. Why do you think this should be added?
6. In what order do you think we should present the sections of privacy, security, and general information to consumers? Why?

D.2 Questions Asked from Privacy and Security Experts on the First Survey

Please answer the following questions and provide us with the reasons that are most convincing to you. The reasons can be in your own words or alternatively, you can quote any resource that you think provides a well-explained reason. Please note that to design the privacy and security label, we will follow a layered approach. In a layered design, we will put the most important/understandable information on the first/primary layer and the additional important information will go on the secondary layer in case consumers are curious to know more or have privacy and security expertise and are looking for more information (the secondary layer can get updated if the value of any factor is changed). Therefore, you may believe that it is important to include a specific factor somewhere on the label, but not necessarily on the first/primary layer. Please also do not worry about the physical space on the label and how many factors should be included on it.

1. For each factor, please select the statement that you believe is correct.
 - Is an important factor to include on the label.
 - Is not an important factor to include on the label.
 - May or may not be an important factor to include on the label.
2. (If the answer to question 1 is “Is an important factor to include on the label.” or “May or may not be an important factor to include on the label.”) Please specify your most convincing reason(s) as to why this is an important factor to put on the label. [text entry]
3. (If the answer to question 1 is “Is not an important factor to include on the label.”) Please specify your most convincing reason(s) as to why this is not an important factor to put on the label. [text entry]

D.3 Questions Asked from Privacy and Security Experts on the Second Survey

For each factor, please read the arguments for and against having that factor on the label (primary or secondary layer) and answer the questions. We selected the factors as well as the arguments from the experts we interviewed.

The label is not only to educate consumers to make informed purchase decisions, but also to hold IoT companies accountable for their privacy and security practices. In addition, all of the factors need to have a precise definition and metric, but for now do not worry about that and assume that the definition and metric is specified.

Including a factor on the label that describes **whether or not the device is getting cryptographically signed and critical automatic security updates**:

- **Is important** because:

- Consumers understand the importance of automatic security updates.
- New vulnerabilities in software are discovered very often and hence continuous checking for updates is needed.

- **In not important** because:

- Signed security updates do not guarantee a well implemented and trustworthy update process.
- Sometimes there are reasons not to have automatic security updates. These reasons include legal considerations or if the device is incapable of receiving updates.
- Automatic updates should be a requirement and there is no need to explicitly mention them.

1. I believe this factor:

- Should definitely be on the label
- Should probably be on the label
- I have no preference about this factor
- Should probably not be on the label
- Should definitely not be on the label

2. On which layer of the label do you think this factor should be presented?

- Primary/first layer (Please specify your reason(s)) [text entry]
- Secondary/extended layer (Please specify your reason(s)) [text entry]
- I have no preference over this (Please specify your reason(s)) [text entry]

3. Information about this factor is for (select all that apply):

- Consumers' understanding
- Experts' understanding
- For holding manufacturers accountable, e.g., using regulations
- Other (Please specify) [Text entry]

4. If the label had separate security, privacy, and general sections, this factor would be most appropriate for:

- The security section
- The privacy section
- Either the privacy or the security section
- The general information section
- None (Please specify) [text entry]

5. Please provide any additional arguments you may have for or against including this factor on the label. [Text entry]

6. Please provide any comments you have about this factor. [Text entry]
7. I believe:
 - Privacy factors and security factors should be presented separately on the label (Please specify your reason(s)) [Text entry]
 - Privacy and security factors should be combined and shown on the label without separation (Please specify your reason(s)) [Text entry]
 - I have no preference toward presenting privacy and security factors (Please specify your reason(s)) [Text entry]

D.4 Interview Questions with IoT Consumers

[first without the label, then with the label] I want you to please take a look at this smart security camera package. Please let me know what can you learn about the privacy and security of this device?

D.4.1 Questions on Risk Communication in Comparative Purchase Process

Let's say you were trying to decide whether to buy the security camera that I just showed you. The store has this other security camera, which is the same price and has the same features. However, it has slightly different privacy and security related information, which is now being presented on the label on the package. Please take a look at these two devices and tell me

1. What are the things that this company has done better over the other one and why do you think they are better?
2. What are the things that this company has done worse and why do you think they are worse?
3. Overall, which device would you purchase? Why?
4. What is the most useful piece of information on this label that helped you make this decision?

[explain the idea of a layered label] Please scan the QR code or type in the URL on the primary layer to take a look at the secondary layer of the label.

1. Except the devices that I showed you in our study, have you ever seen a product that has a layered label like we describe?
2. What is your opinion on the idea of a layered label?
3. What could be the advantages and disadvantages of a layered label as opposed to a simple label that does not have layers?

D.4.2 Questions on Information Comparison in Non-Comparative Purchase Process

1. For each presented factor in the security mechanisms/privacy practices/general information section, can you please explain what the factor means?
2. How useful do you believe this factor would be in your purchase decision making?
3. Do you have any suggestions to make the information about this factor more understandable?

D.4.3 Questions on Risk Communication in Non-Comparative Purchase Process

1. What are the factors that you see here that seem risky to you from a security or privacy perspective?
2. (If participants specify a factor as being risky) What kinds of risks do you think you could be exposed to by the information conveyed by this factor? What bad things could happen as a result of this information?
3. If you were looking for a camera that was like this but less risky, what would you like to see instead of the current information for this factor?
4. If features, price, and the brand were all what you wanted, based on the privacy and security information of this primary layer of the label, would you purchase this device? Why?
5. What are the most useful factors on this label that helped you make this decision?

D.4.4 Questions on Information Comparison and Risk Communication of the Secondary Layer

1. For each presented factor in the security mechanisms/privacy practices/general information section, can you please explain what the factor means?
2. How useful do you believe this factor would be in your purchase decision making?
3. Do you have any suggestions to make this information more understandable?

D.4.5 Questions on Format and Layout Considerations

1. Overall, what do you think about the current design consideration of separating out security mechanisms, privacy practices, and general information?
2. What are the changes you would like to apply to this label to make it more usable and understandable to you and other consumers? The suggestions could be both related to the content of the label or the design of the label. Please specify your reasons for the change you proposed.
3. Is there anything else that you think would be useful to have on the primary layer of the label?

4. Is there anything else that you think would be useful to have on the secondary layer of the label?

D.4.6 Questions on Purchase Behavior Scenarios

1. If you were going to purchase a smart device from a physical store, where would you go?
2. If you were going to purchase a smart device from an online store, where would you go?

Which one is the kind of process that you would be most likely to follow if you were to purchase a smart home device at brick & mortar? Why? When reading the scenarios, please replace the Home Depot, with [physical store participants mentioned] and the online store with [online store participants mentioned].

- You walk into HomeDepot and go to the smart home aisle. You find the products you are interested in and see the primary layers of the labels on the packages. You stand there looking at the packages, but you don't go online to take a look at the secondary layer.
- You walk into HomeDepot and go to the smart home aisle. You find the products you are interested in and see a QR code/URL on the package. You stand there and scan the QR code/type in the URL and look up the information of the secondary layer on your phone.
- Before going to HomeDepot, you go online and do the comparison shopping. You come to HomeDepot already knowing which device you will buy.
- You do the comparison shopping online, find the product you are interested in and purchase the device from Amazon.
- Something else? Please explain.

Appendix E

Survey Questions from “Which Privacy and Security Attributes...”

E.1 Survey Questions

Questions presented here are for one of the experimental conditions, in which the between-subject factors of *device type* and *device recipient* are smart light bulb and purchasing the device for a friend.

E.1.1 Device-Related Questions

1. How concerned are you about the way *smart light bulbs* with *presence sensors* collect, store, and use information related to *whether someone is present in the room*?
 Not at all concerned
 Only slightly concerned
 Somewhat concerned
 Moderately concerned
 Very concerned
2. (If the answer to question 1 is “Only slightly concerned,” “Somewhat concerned,” “Moderately concerned,” or “Very concerned”) What about data collection, storage, and use by *smart light bulbs* with *presence sensors* makes you feel concerned? [text entry]
3. (If the answer to question 1 is “Not at all concerned”) What about data collection, storage, and use by *smart light bulbs* with *presence sensors* makes you feel not at all concerned? [text entry]
4. Do you currently have a *smart light bulb* with *presence sensor* in your home?
 Yes
 No
5. (If the answer to question 4 is “Yes”) How long have you had your *smart light bulb* with *presence sensor*? If you have more than one device, answer the question for the one that you have had for the longest time.

- Less than a month
 - Between a month and a year
 - More than a year
 - I don't remember
6. (If the answer to question 4 is "Yes") How did you acquire your *smart light bulb* with *presence sensor*? (check as many as apply)
 - I purchased it
 - Somebody else in my home purchased it
 - I received it as a gift
 - It was installed by my landlord
 - Other (please specify) [text entry]
 7. (If the answer to question 6 is "I purchased it") What brand(s) of *smart light bulb* with *presence sensor* did you purchase? [text entry]
 8. (If the answer to question 6 is "I purchased it") What were your reasons to purchase a *smart light bulb* with *presence sensor*? [text entry]
 9. (If the answer to question 4 is "No") Have you ever been in the market to purchase a *smart light bulb* with *presence sensor*?
 - Yes
 - No
 10. (If the answer to question 9 is "Yes") What made you decide not to purchase the *smart light bulb* with *presence sensor* that you were in the market for? [text entry]

E.1.2 Label-Related Questions

Here we only present the questions that participants will see if one of the randomly-assigned *attribute-value* pairs is *security update-automatic*.

Imagine you are making a decision to purchase a *smart light bulb* for a *friend*. This device has a *presence sensor* that will *detect whether someone is present in the room to control the lighting automatically*. The price of the device is within your budget and the features are all what you would expect from a smart light bulb with presence sensor. On the package of the device there is a label that explains the privacy and security practices of the *smart light bulb with presence sensor*.

The label on the device indicates the following:

Security update: Automatic

1. How confident are you that you know what this information on the label means?
 - Not at all confident
 - Only slightly confident
 - Somewhat confident
 - Moderately confident
 - Very confident

Risk Perception

1. I believe receiving *automatic security updates*
 - Strongly decreases the privacy and security risks associated with this specific *smart light bulb with presence sensor*
 - Slightly decreases the privacy and security risks associated with this specific *smart light bulb with presence sensor*
 - Does not have any impact on the privacy and security risks associated with this specific *smart light bulb with presence sensor*
 - Slightly increases the privacy and security risks associated with this specific *smart light bulb with presence sensor*
 - Strongly increases the privacy and security risks associated with this specific *smart light bulb with presence sensor*
 - I am not sure how *receiving automatic security updates* impacts the privacy and security risks associated with this specific *smart light bulb with presence sensor*
2. (If the answer to question 1 is “Strongly decreases the privacy and security risks associated with this specific *smart light bulb with presence sensor*” or “Slightly decreases the privacy and security risks associated with this specific *smart light bulb with presence sensor*”) Please explain why you believe *receiving automatic security updates* decreases the privacy and security risks associated with this specific *smart light bulb with presence sensor*. [text entry]
3. (If the answer to question 1 is “Strongly increases the privacy and security risks associated with this specific *smart light bulb with presence sensor*” or “Slightly increases the privacy and security risks associated with this specific *smart light bulb with presence sensor*”) Please explain why you believe *receiving automatic security updates* increases the privacy and security risks associated with this specific *smart light bulb with presence sensor*. [text entry]
4. (If the answer to question 1 is “Does not have any impact on the privacy and security risks associated with this specific *smart light bulb with presence sensor*”) Please explain why you believe *receiving automatic security updates* does not have any impact on the privacy and security risks associated with this specific *smart light bulb with presence sensor*. [text entry]

Willingness to Purchase

1. Assuming you were in the market to purchase this *smart light bulb with presence sensor* for *a friend as a gift*, knowing that the device will *automatically receive security updates*, would
 - Strongly decrease your willingness to purchase this device for *a friend as a gift*
 - Slightly decrease your willingness to purchase this device for *a friend as a gift*
 - Not have any impact on your willingness to purchase this device for *a friend as a gift*
 - Slightly increase your willingness to purchase this device for *a friend as a gift*
 - Strongly increase your willingness to purchase this device for *a friend as a gift*

2. (If the answer to question 1 is “Strongly decrease your willingness to purchase this device for a friend as a gift” or “Slightly decrease your willingness to purchase this device for a friend as a gift”) Please explain why knowing that the device will *automatically receive security updates* would decrease your willingness to purchase the device for a friend as a gift. [text entry]
3. (If the answer to question 1 is “Strongly increase your willingness to purchase this device for a friend as a gift” or “Slightly increase your willingness to purchase this device for a friend as a gift”) Please explain why knowing that the device will *automatically receive security updates* would increase your willingness to purchase the device for a friend as a gift. [text entry]
4. (If the answer to question 1 is “Not have any impact on your willingness to purchase this device for a friend as a gift”) Please explain why knowing that the device will *automatically receive security updates* would not have any impact on your willingness to purchase the device for a friend as a gift. [text entry]

Attention Check Question for Automatic Update

1. Which statement is correct about the device described in the previous question?
 - The device will automatically get updated
 - The device will manually get updated
 - The device will not get updated
 - The device will ask for my permission each time to install security updates

E.1.3 Additional Attributes

The complete list of privacy and security attributes specified by Emami-Naeini et al.’s [114] are grouped into three sections: *general*, *privacy*, and *security*. For each section, we present participants with a table, which rows are the attributes specific to that section and the columns are “Interested to know about” and “Would impact my willingness to purchase”. For instance for the security section the question would be:

If you had the opportunity to know more about the *security practices* of the *smart light bulb* with *presence sensor*, which of the practices below would you be more interested to know about and which would impact your willingness to purchase the device for a friend as a gift? (check as many as apply)

One of the rows is: “Safeguards the manufacturer has in place to protect the device hardware from being tampered with.” and participants can specify whether they are interested to know about this information and whether this information would impact their willingness to purchase the device.

E.1.4 Functionality Perception

1. How do you think a *smart light bulb* with *presence sensor* works?
 - The device always senses whether someone is present in the room
 - The device starts sensing whether someone is present in the room when you press a button to turn on the presence sensor on the device
 - The device starts sensing whether someone is present in the room when you turn on the lights
 - I have no idea how a smart light bulb with presence sensor works

E.1.5 Demographic Questions

1. What is your age? [text entry]
2. What is your gender? [text entry]
3. What is the highest degree you have earned?
 - No high school degree
 - High school degree
 - College degree
 - Professional degree (Master's/PhD/medical/law)
 - Associate degree
 - Prefer not to answer
4. Do you have a background in technology (if yes, please specify what your background is)? [text entry]

E.1.6 Consumer Explanations for Attribute-Value Pairs

Attribute-value	Consumer explanation
Security update: automatic	Device will automatically receive security updates
Security update: none	Device will not receive any security updates
Access control: multi-factor authentication	At least two independent factors to authenticate a user are required to access the device, for example a password and a confirmation from a previously registered phone
Access control: none	Anyone can access the device without a password or other authentication method
Purpose: device function	Data is being collected to provide the main device features, improve services, and help develop new features
Purpose: monetization	The manufacturer and service provider receive income from showing personalized advertisements to users or selling user's data to third parties
Device storage: none	The collected data will not be stored on the device
Device storage: identified	User's identity could be revealed from the data stored on the device
Cloud storage: none	The collected data will not be stored on the cloud
Cloud storage: identified	User's identity could be revealed from the data stored on the cloud
Shared with: none	Data is not being shared
Shared with: third parties	Data is being shared with third parties
Sold to: none	Data is not being sold
Sold to: third parties	Data is being sold to third parties
Average time to patch: 1 month	Vulnerabilities will be patched within 1 month of discovery
Average time to patch: 6 months	Vulnerabilities will be patched within 6 months of discovery
Security audit: internal & external	Security audits are performed by internal and third-party security auditors
Security audit: none	No security audit is being conducted
Collection frequency: on user demand	Data is collected when the user requests it
Collection frequency: continuous	When the device is turned on, it will continuously collect data until it is turned off
Sharing frequency: on user demand	Data is shared when the user requests it
Sharing frequency: continuous	When the device is turned on, it will continuously share data until it is turned off
Device retention: none	User's data will not be retained on the device
Device retention: indefinite	User's data may be retained on the device indefinitely
Cloud retention: none	User's data will not be retained on the cloud
Cloud retention: indefinite	User's data may be retained on the cloud indefinitely
Data linkage: none	Data will not be linked with other data sources
Data linkage: internal & external	Data may be linked with other information collected by the manufacturer as well as other information
Inference: none	No additional information about user will be inferred from user's data
Inference: additional information	User's characteristics and psychological traits, attitudes and preferences, aptitudes and abilities, and behaviors could be inferred from the collected data
Control over: cloud data deletion	User has an option to delete the data that is being stored on the cloud
Control over: device storage	Data will not be stored on the device unless user opts in to device storage
Control over: device retention	User can change the duration for which their data may be retained on the device

Table E.1: In the survey, we provided a consumer explanation for each attribute-value pair.