Designing Privacy Notices
Supporting User Understanding and Control

Patrick Gage Kelley

May 2013
CMU-ISR-13-106

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee
Lorrie Faith Cranor, co-chair
Norman Sadeh, co-chair
Alessandro Acquisti
Sunny Consolvo

Submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy

This research was supported by the National Science Foundation under grant numbers CNS-1012763, CNS-0905562, CNS-0627513, CNS-0831428, and DGE-0903659, and by the U.S. Army Research Office grants DAAD19-02-1-0389 and W911NF-09-1-0273.
Additional support was provided by Carnegie Mellon University's CyLab, Intel Labs Seattle, the University of Washington, Google, Microsoft through the Carnegie Mellon Center for Computational Thinking, FCT through the CMU/Portugal Information and Communication Technologies Institute, and the IBM OCR project on Privacy and Security Policy Management.

Keywords: privacy, notice, usability, user interfaces, security, mobile, policy, P3P, HCI, information design

Designing Privacy Notices

Supporting User Understanding and Control

Patrick Gage Kelley

Doctorate of Philosophy in Computation, Organizations & Society
CyLab Usable Privacy and Security Doctoral Training Program
Institute for Software Research
School of Computer Science
Carnegie Mellon University

COMMITTEE

Lorrie Faith Cranor
Institute for Software Research & Engineering and Public Policy

Norman M. Sadeh
School of Computer Science

Alessandro Acquisti
Heinz College

Sunny Consolvo
Google

Designing Privacy Notices

Text set in Chaparral, a slab serif typeface designed by Carol Twombly.

Design by Patrick Gage Kelley

www.patrickgage.com

Computation, Organizations & Society
CyLab Usable Privacy and Security Doctoral Training Program
Institute for Software Research
School of Computer Science
Carnegie Mellon University

DEDICATION

to the teachers...

Diane Kelley

George Kelley

Norman Schoell

Laurence Hiller

Jessica Lieberman

Amit Ray

ABSTRACT

Users are increasingly expected to manage complex privacy settings in their normal online interactions. From shopping to social networks, users make decisions about sharing their personal information with corporations and contacts, frequently with little assistance. Current solutions require consumers to read long documents or go out of their way to manage complex settings buried deep in management interfaces, all of which lead to little or no actual control.

The goal of this work is to help people cope with the shifting privacy landscape. While our work looks at many aspects of how users make decisions regarding their privacy, this dissertation focuses on two specific areas: the current state of web privacy policies and mobile phone application permissions. We explored consumers' current understandings of privacy in these domains, and then used that knowledge to iteratively design and test more comprehensible information displays.

These prototyped information displays should not be seen as final commercially-ready solutions, but as examples of privacy notices that can help users think about, cope with, and make decisions regarding their data privacy. We conclude with a series of design suggestions motivated by our findings.

CONTENTS

Smartphone Application Privacy

Conclusion

LIST OF FIGURES

LIST OF TABLES

7.5 Timing information for the application-selection task across both conditions and the time spent on permissions screens by participants viewing the permissions display. Participants in the Privacy Facts condition could not look at the standard android permissions displays. The Number of seconds spent on permissions column shows how long a participant looked at a permissions screen each time they did so. 115

7.6 Application selections in the laboratory and online studies. The application that requested access to fewer permissions (the privacy-protective choice) is always displayed on top. Statistics for the online study are comparisons to the base Android display. The right-most column shows the significance between the checklist and the inline permissions. Differences in bold, Fisher's Exact. Comparisons with the Android display were planned contrasts. The final comparison between the permissions inline and privacy facts display is Holm-corrected with an adjusted alpha of 0.01667. 116

INTRODUCTION

1

1.1 MOTIVATION

Alan Westin defines one aspect of privacy in his seminal 1967 *Privacy and Freedom* as a person "continually engaged in a personal adjustment process in which he balances the desire for privacy with the desire for disclosure and communication of himself to others" [112, p. 7].

The idea of representing privacy as an on-going process is one that still rings true 45 years later. And across the last 45 years, we have seen a rate of technological progress that has forced people to negotiate this desire for privacy in new and unexpected ways.

Westin wrote *Privacy and Freedom* with electronic databases just over the horizon. Now databases are ubiquitous and the internet has become a world of electronic commerce, social networking, and tracking. The flow of users' information has reached a new and unprecedented era.

In August 2012, Facebook reported they receive 300 million photos each day, processing 105 terabytes of data every half hour [19]. Each of these images are posted with user-specified access restrictions, may be linked to other users through "tags," and may be geo-tagged. Sharing photos has changed over the last two decades from gathering around a physical album in one's living room to managing this array of access and metadata controls online. This spread of managing personal information access does not only apply to photos. Looking just at geo-tagged information, people now "check-in" to location-based social networks. Foursquare, the current genre leader, has 3 billion checkins, adding millions each day [44].

Many of these activities are made possible by the ever increasing use of smartphones. Over 900 million people globally have active smartphone subcriptions [103]. Having a personal computer in your pocket affords apps access to a device that is always with a user and always on. These apps allow new types of information sharing, but require users to install software. The risks of granting this access to

developers across the globe and methods for mitigating access to the vast troves of personal data contained within are still coming to light.

In an infographic and article in May 2010 the New York Times attempted to understand one facet of the data management costs which are now part of privacy. Documenting the over 170 options over 50 different settings that Facebook currently offers users, Gates said, "privacy groups, government officials and its own users, ... complain that the new policy is bewildering and the new opt-out settings [are] too time-consuming to figure out and use" [13, 46]. The current state of privacy controls on Facebook leaves users unable to make informed decisions. Yet, users are making these decisions every day with unknown consequences, providing their information to a company with a privacy policy longer than the United States Constitution.

Privacy policies have been shown to not provide an adequate level of notice and consent. Most policies are written at a level that is suitable for consumers with a college-level education and use specific domain terminology that consumers are frequently unfamiliar with [55, 78]. Furthermore, even when the policies may be comprehensible, reading current online privacy policies is time consuming. It is estimated that if every internet user read the privacy policies for each site they visited, this lost time would cost about $365 billion per year [77]. Yet we see from e-commerce to location sharing, sites frequently fall back on privacy policies as the primary form of communication to users about their data practices [107].

The launch of Google Buzz and it's automatic contact setup led to a user backlash [82], immediate product alterations, and eventually the establishment of an $8.5 million dollar privacy fund [15]. The Library of Congress set off media alarms when it announced it would archive all of Twitter, including users' past tweets [87]. Most recently, the internet whined after Instagram, a popular photo-sharing mobile app that had been recently acquired by Facebook, changed it's Terms of Service [14]. The change caused users to fear their photos would be misappropriated, leading many to threaten to quit the service.

Better designed and implemented privacy interfaces and notices could have solved or alleviated the end-user concerns in many of the above stories. A better designed contact import tool (which allowed users to opt-in) could have saved Google Buzz's initial rollout. More awareness and explanation of the "protect my tweets," functionality Twitter provides would have allowed users a simple way to opt-out of archiving by the Library of Congress. Had Instagram initially

explained the rationale for their Terms of Service changes, the fears raised by consumers and the media would likely have never occurred.

These high profile examples show that there is a continued disconnect between users' expectations of privacy and the data practices and settings of the services they use.

There is not a single solution for this problem nor is it possible to be solved entirely. As Westin notes [112, p. 7], people balancing their privacy desires is a process, and as technology continues to improve and evolve consumers will be faced with different privacy management interfaces and platforms and changing expectations. However, consumer education, better information and awareness of corporate practices, and protective legislation can all serve to improve the current state of privacy.

1.2 RESEARCH NARRATIVE

This work was guided by many of the examples above, and other similar instances of consumers being unable to reason and make decisions about the data practices of companies they do business with. To bridge this gap we sought to focus on privacy notices, given the central role of notice in the Fair Information Principles [34]. Specifically:

Thesis Statement

The goal of this work is to explore how improved privacy notices can be created and iteratively improved to help consumers better understand data practices and take more active control of their information.

The thesis explores the creation of improved privacy notices in two domains: online privacy policies through a "nutrition label" approach and smartphone application selection through short permission notices at the point of decision.

The first domain we explored was online privacy policies. Growing organically from work at the CyLab Usable Privacy and Security (CUPS) Lab, we developed a series of displays of online privacy policies. We based these designs off of the Platform for Privacy Preferences (P3P), a standardized XML-based format for representing privacy policies. These designs were partially a response to the ongoing research assessment of online privacy policies as a

consumer failure. We tested our policy format for understanding, time to use, ability to facilitate comparisons, and user enjoyment.

Our prototyping and large-scale, though not real-world, testing shows that the "nutrition label" format for privacy policies is better than the current market offerings. Our format is standardized and informed by user comprehension and testing. While our design performed well in these tests, due to sparse P3P adoption, either legislation or a widespread corporate shift would need to take place for the label to be widely deployed.

As such, we moved to a much more competitive market, that of smartphone applications. Here consumers are making more decisions, have real-time information at stake, and are making a portion of their decisions with no brand effects (downloading from previously unknown developers). Specifically, we began to explore the consumer response to the application permissions display on Android smartphones. By leveraging the way consumers currently make decisions we developed a short privacy notice to better inform consumers about applications' data practices at the point of their decision.

1.3 ORGANIZATION

Chapter 2 will provide an introduction to the current state of online privacy policies. Additionally, it will provide a brief overview of research towards improving privacy policies as well as work in consumer labeling and education around nutrition and financial privacy. Chapter 3 details the design work and early testing of our nutrition label for online privacy, while Chapter 4 documents the large-scale verification of the design.

The state of Android smartphone research is summarized in Chapter 5. Chapter 6 details our exploratory interview study of smartphone users' comprehension of the Android permission model. Chapter 7 describes our process for designing an Android privacy notice, and provides the results of an in-person and an online experiment testing the notice.

Finally, Chapter 8 highlights lessons learned from across the two domains and concludes with several specific design patterns and trends that we witnessed throughout our experiments.

PART I

ONLINE PRIVACY POLICIES

PRIVACY POLICIES AND CONSUMER UNDERSTANDING

RELATED WORK

2

Privacy policies fail to provide consumers with even a basic level of notice. Here we look at a brief history of how they came to be, at the literature on why and how they are unsuccessful, and lessons learned from other standardization efforts on consumer notice.

2.1 A BRIEF HISTORY OF PRIVACY POLICIES

Fifteen years ago the internet was unrecognizably different. Not just because web design was a fledgling area brimming with table-based layouts and javascript had not yet been optimized and exploited to provide the responsiveness we now expect, but because companies did not provide any notice about their business practices. The now common checkbox affirming that you agree with a site's Terms of Service and Privacy Policy were next to nonexistent, as were links in the footer of every page providing access to that same information.

From the Federal Trade Commission's (FTC) June 1998 "Privacy Online: A Report to Congress" they found that in their own survey of over 1400 websites, over 85% were collecting personal information from consumers [35]. However only 14% of those sites were providing any sort of notice regarding what was done with that personal data and only 2% provided a comprehensive privacy policy. That means of the most popular sites on the internet only 2%, or around 30 websites, provided the now ubiquitous "Privacy" link.

That report continues, "The Commission has encouraged industry to address consumer concerns regarding online privacy through self-regulation. The internet is a rapidly changing marketplace. Effective self-regulation remains desirable ... To date, however, the Commission has not seen an effective self-regulatory system emerge." This means that the FTC believed these small percentages of notice were unsatisfactory. This is completely in line with one of the FTC's goals, to support the commonalities between fair information practice codes, also described as the "five core principles of privacy protection:"

1. Notice/Awareness
2. Choice/Consent
3. Access/Participation
4. Integrity/Security
5. Enforcement/Redress

While, we will not describe each of these in detail, we highlight notice, as it is the core tenant of the remainder of the thesis. The FTC introduces notice as follows [34]:

> **"The most fundamental principle is notice.** Consumers should be given notice of an entity's information practices before any personal information is collected from them. Without notice, a consumer cannot make an informed decision as to whether and to what extent to disclose personal information. Moreover, three of the other principles discussed below are only meaningful when a consumer has notice of an entity's policies, and his or her rights with respect thereto."

Notice is providing consumers with information about how an entity uses, collects, stores, and shares information. The idea here is that this information is mandatory to help consumers make smart decisions regarding which entities they trust with their personal information.

The FTC Chairman Robert Pitofsky concluded the report by recommending that Congress pass legislation if self-regulation failed to produce significant progress. However, by 1999 comprehensive privacy policies were found on over 80% of the then top websites.

With this remarkable turn around in notice companies could claim that they had successfully self-regulated, providing consumers with notice.

However, while websites were lacking in 1998 because they provided no privacy information, we argue the mere existence of such notices is not enough. The following years have shown that these notices are inadequate, with little response or change from companies.

2.2 THE UBIQUITY OF PRIVACY POLICIES

"Industry progress has been far too slow since the Commission first began encouraging the adoption of voluntary fair information practices in 1996. Notice, while an essential first step, is not enough if the privacy practices themselves are toothless."

– Commissioner Sheila Anthony [5]

Internet privacy is largely unregulated in the United States, with only a few exceptions for the protection of children's privacy and some sector-specific regulations. As such companies are free to write their privacy policies as they see fit. While the FTC did begin charging companies for omission and misrepresentation of deceptively collecting personal information at this time, even in these cases the guidance for better notice remained vague. In this case the FTC "require[d] the company to post on its site a clear and prominent Privacy Notice, telling consumers what information is being collected and for what purpose, to whom it will be disclosed, and how consumers can access and remove the information" [33]. This has led to a number of specific failings being aimed at today's privacy policies. Cranor has written more on the sufficiency of notice regarding today's privacy policies [20].

2.2.1 Awareness

It has now been established through numerous studies that people do not read privacy policies [89] and make mistaken assumptions based upon seeing that a site has a link to a privacy policy [109].

Even worse, it is unlikely consumers will even read a single policy given a widespread consumer belief that there are no choices when it comes to privacy. Many consumers believe they do not have the ability to limit or control companies' use of their information [64]. This is a finding that we again validated in our work.

2.2.2 *Time*

Even if consumers believed there were a reason to read privacy policies, and decided to do so, they would find that reading privacy policies is a time consuming activity. It has been estimated that if every internet user read the privacy policies for each site they visited, this time would cost about $365 billion per year in lost productivity [77]. This amount of time, 244 hours, per year, per person, is greater than the time to handle spam, and on par with the current time spent surfing the web. While the idea of a person reading this many privacy policies is admittedly unrealistic, it provides a benchmark to estimate one failing of today's privacy policies.

2.2.3 *Readability*

A part of the reason that reading even a single privacy policy takes so much time is that any one policy was not created based upon a standardized, learnable format. Most online privacy policies are written by a committee of managers, consultants, and lawyers, and use jargon-heavy, legal language [55, 78]. The result of this is a final document with a readability level that is congruent with a college education. In a study published in 2004, the readability of the top 50 site privacy policies was found to be a 34.2, scoring between the Wall Street Journal and the Harvard Law Review [55][1].

1. This study used Flesch-Kincade Reading Ease. For comparison, The Wall Street journal averages a 43, the Harvard Law Review a 32, where lower scores are more difficult to read

Additionally, policies are often not written to assist a consumer using the website, but are often more reflective of the ways data is understood within a corporate system. Rarely is a policy written such that consumers have a clear understanding of where and when their data is collected, how and by whom it will be used, if it will be shared outside of the entity that collected it, and for how long and in what form it will be stored [60].

In theory, we are painted a picture of perfect industry self regulation, where companies voluntarily post online privacy policies, which are FTC enforced for deceptive and fraudulent practices. Consumers can then visit new sites, read and understand their privacy policies, compare different sites' policies, and then select a site that matches a level of personal sharing they feel comfortable with. And all of this can be done by a consumer in combination with all the other factors that go into e-commerce and online account creation: brand, price, features, and so on. This, again in theory, should produce a market that benefits consumers (through choice) and benefits companies with strong privacy protections (through selection).

In practice, while privacy policies are posted and FTC enforced at some minimal level, consumers visit new sites, and make their decisions without ever considering the posted privacy policies. Today's privacy policies are not providing informed notice. They do not convey an understanding of how companies are using, storing, and sharing personal data, because they are too long, too difficult to read, and habitually ignored.

2.3 STANDARDIZED LABELING PROGRAMS

While explaining corporate privacy information and data practices to consumers is a relatively new area, others have sought to inform consumers about complicated subject matter. To better inform our design process we surveyed the literature these consumer labeling efforts: the "Nutrition Facts" panel, energy and drug labeling, and recent work on creating a standardized financial privacy notice. These are nearly always government led efforts, with external organizations fulfilling the design and testing work.

2.3.1 The "Nutrition Facts" Panel

In the United States, the nutrition label seen in Figure 2.1, has become iconic after being mandated by the Nutrition Labeling and Education Act of 1990 (NLEA) [40].[2] In the last two decades, its increasing ubiquity has led to a number of studies examining the costs of adoption and the ability to inform and change consumer purchasing decisions.

2. Before the passage of this bill, food labeling was adhoc and often full of misleading health claims.

The sparse literature around the design of the nutrition label [12] focuses on the decisions which heavily simplified the information for consumers. These decisions were made in part to address low literacy rates and the needs of older Americans, and will be leveraged throughout this work. These guidelines include equalizing labels across products by providing quantitative information about nutrients, defined serving sizes, and calculating percentages off of standardized daily amounts. They also include the hallmarks of standardization, defining a zone of authority, specifying line weights, selecting typefaces, and defining minimum font sizes.

Surveys indicate that consumers report they would prefer that nutrition labels include more information. However, studies have shown that including more information would not actually be beneficial [27]. Studies conducted to examine the impact of the NLEA have found that it is the populations of people who are educated and already motivated to investigate nutritional

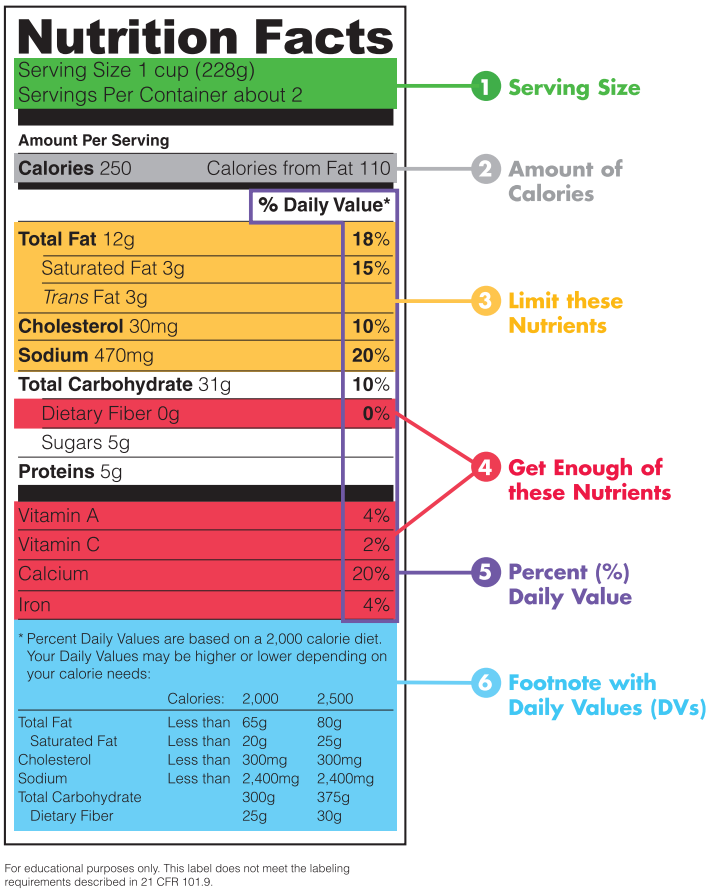


Figure 2.1: The Food and Drug Administration's Nutrition Facts panel as regulated by the NLEA. Source: www.fda.gov.

information who benefit the most from nutrition labels [11, 27]. One study found that nutrition information had the greatest impact when there was a limited number of items from which to make a selection [99]. This result implies that the nutrition label made it easier to compare between a small set of items, allowing consumers to benefit, through informed decision making, when the information presented to them was artificially limited in quantity. Studies have demonstrated that nutrition labels do have an impact on real-world consumer decision making. Yet, even user-reported effect sizes up to 48% after the initiation of NLEA [27], most studies have focused on specific nutrients or products such as fat-intake or salad dressings, instead of universal effects. We are still not aware of longitudinal studies to measure the impact of nutrition labels on consumer behavior over an extended period of time.

Other studies have found that the effects of providing calorie information (not a complete nutrition facts label) in restaurant menus are often very small and the effects may vary depending on the population studied. In a study of meal choices at a sandwich

shop, Downs et al. [26] found that participants given menus that included calorie information ordered meals with about 50 fewer total calories than participants given menus that did not include calorie information. However, the authors stated that this was "an effect smaller than this study was powered to test." Nonetheless, they pointed out that if the finding proved reliable, it could be significant if it caused people to reduce their caloric intake by a similar amount for multiple meals each day. In a related study of food purchases at three New York City restaurants before and after a law went into effect mandating the posting of calorie information on menu boards, the authors found no effects of the legislation at two of the three restaurants. At the third restaurant they found a small effect. They noted that the effect was larger for dieters than for non-dieters, suggesting again that the availability of label information may be most useful to people who are already interested in the information provided by the label [26].

2.3.2 Other Labeling Programs

We also explored energy labeling programs from the European Union [18] and Australia [85], the US Consumer Products Safety Commission's toy and game warnings [23], and the US FDA Drug Facts label [42], to gain a broader understanding of practices used in designing and defining labeling requirements.

In general, the standards documents [17, 18, 40] are occupied with defining precise guidelines to describe compliance with the various labeling requirements. This includes point sizes of rules and text, allowable typefaces, allowable colors, and minimum sizes. In some instances, such as choking warnings on children's games, standards also include placement requirements.

Recently, a number of labels have been introduced to provide ratings to consumers on a fixed scale, focusing on a single metric or a small number of metrics. The Australian Water Efficiency Labeling System (WELS) [50] and the British Food Standards Agency's Signposting (or Traffic Light) [2] use very small indicators with accompanying ratings. The WELS program uses an indicator with a possible score out of six blue stars, Figure 2.2. The Signposting initiative rates the quantities of fat, saturates, sugar, and salt in foods using a red, amber, green traffic light coloring system, Figure 2.3. Early research [11, 75] has shown that Signposting enhances consumers ability to evaluate products more accurately and surveys show ninety percent of consumers find this type of label useful.

Figure 2.2: Example of the Australian Water Efficiency Labeling System (WELS), retrieved April 2013, from http://www.waterrating.gov.au/

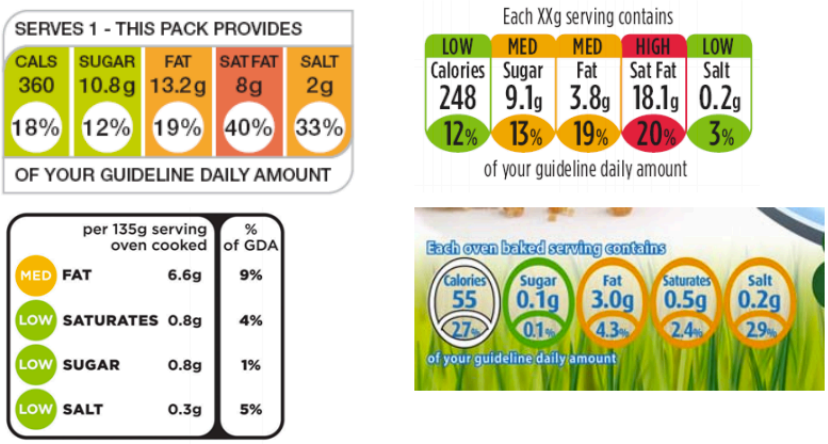


Figure 2.3: Examples of labels in the marketplace with combined traffic light and daily value percentages. From BMRB Report: *Comprehension and use of UK nutrition signpost labelling schemes*. Retrieved April 2013, from http://www.food.gov.uk/multimedia/pdfs/pmpreport.pdf

2.3.3 *Financial Privacy Notices*

In 2004, seven federal agencies[3] launched a multi-phase initiative to "explore the development of paper-based, alternative financial privacy notices...that are easier for consumers to understand and use" [64]. That is, their goal was similar to our goal in the next two chapters, but was focused around financial privacy specifically.

3. The seven federal agencies that enforce the GLBA are the Federal Deposit Insurance Corporation, the Federal Reserve Board, the Federal Trade Commission, the National Credit Union Administration, the Office of the Comptroller of the Currency, the Office of Thrift Supervision, and the Securities and Exchange Commission.

The Kleimann Communication Group (KCG) used an iterative design process to develop a prototype notice, focusing on user comprehension, allowing users to "identify differences in sharing practices," and compliance with the regulations surrounding financial privacy notices specified in the Gramm-Leach-Bliely Act. The Gramm-Leach-Bliley Act (GLBA), passed in 1999, contains the Financial Privacy Rule, which requires that financial institutions disclose their privacy policy "at the time of establishing a consumer relationship...and not less than annually" [110]. Financial institutions must comply with requirements on what they disclose, but their disclosures may be in any format.

Over a 12-month period the KCG iterated on several design prototypes, conducting focus groups and diagnostic usability testing [64]. Our iterative design approach followed a similar process of testing labels for comprehension and then overall design through focus groups.

The KCG final prototype is a three page document that consists of four parts, the title, the frame, the disclosure table, and the opt-out form. The disclosure table, which actually displays the company's

F A C T S — **WHAT DOES NEPTUNE BANK DO WITH YOUR PERSONAL INFORMATION?**

Why?

Financial companies choose how they share your personal information. Federal law gives consumers the right to limit some but not all sharing. Federal law also requires us to tell you how we collect, share, and protect your personal information. Please read this notice carefully to understand what we do.

What?

The types of personal information we collect and share depend on the product or service you have with us. This information can include:

- social security number and income
- account balances and payment history
- credit history and credit scores

When you close your account, we continue to share information about you according to our policies.

How?

All financial companies need to share customers' personal information to run their everyday business—to process transactions, maintain customer accounts, and report to credit bureaus. In the section below, we list the reasons financial companies can share their customers' personal information; the reasons Neptune Bank chooses to share; and whether you can limit this sharing.

Reasons we can share your personal information	Does Neptune Bank share?	Can you limit this sharing?
For our everyday business purposes— to process your transactions, maintain your account, and report to credit bureaus	Yes	No
For our marketing purposes— to offer our products and services to you	Yes	No
For joint marketing with other financial companies	Yes	No
For our affiliates' everyday business purposes— information about your transactions and experiences	Yes	No
For our affiliates' everyday business purposes— information about your creditworthiness	Yes	Yes *(Check your choices, p.3)*
For our affiliates to market to you	Yes	Yes *(Check your choices, p.3)*
For nonaffiliates to market to you	Yes	Yes *(Check your choices, p.3)*

Contact Us — Call 1-800-898-9698 or go to www.neptunebank.com/privacy

Figure 2.4: The title, frame, and disclosure table from an example prototype financial notice by the Kleimann Communication Group. From their 2006 Report: *Evolution of a Prototype Financial Privacy Notice.* Retrieved April 2013, from http://ftc.gov/privacy/privacyinitiatives/ftcfinalreport060228.pdf

privacy practices, makes up the majority of our label, Figure 2.4. The rest of the KCG prototype was educational information to build a foundation of terms and understanding for the user [64].

In December 2008, the second phase report was published by Levy and Hastak [70]. This report detailed a 1032-participant mail/interview study that tested four privacy notice formats for three fictional financial institutions.

Two of the four prototypes they tested, were developed by the KCG, with contextual information and an opt-out form. The KCG table notice displayed financial institutions' practices in a grid format,

whereas their prose notice used a bulleted list. The other two notices were both text-based, with the "current notice" mimicking notices that financial institutions currently use, and the "sample clause" notice generated from GLBA-provided phrases.

Levy and Hastak concluded that the KCG table notice performed best "on a diverse set of ... measures." They attributed this improvement to an increased level of comprehension, given the table notice's "[provision] of a fuller context...the part-to-whole display approach seems to help consumers focus on information sharing as important and differentiating features of financial institutions." However, on several study questions other notices, notably the sample clause notice, tested best.

Model forms were distributed after the 2008 report and has since been widely adopted by financial institutions.

2.3.4 Layered Policy Notices

Layered privacy notices, popularized by the law firm Hunton & Williams [104, 105], provide users with a high-level summary of a privacy policy. The design is intended to be a "standardized" format; however, the only standard components are a tabular page layout and mandatory section header text. Other design details and the full text of each section are left to the discretion of each company.

Additionally, the amount of information to include in a layered notice is left up to each company, with layered notices requiring consumers to click through to the standard privacy policy for complete information. This means that a layered policy can give users the information they want or it can act as a roadblock in their way to finding that information, and this is up to the discretion of the company that creates it. We will return to layered policies in our large-scale test in Chapter 4.

2.3.5 Web-based Privacy Notices and Ratings

While web privacy seals (small image badges used to represent everything from privacy protection to encryption to "verified security") and lock icons in browsers and on websites have existed since the early days of online commerce, more complete, visual privacy notices are relatively new.

Companies like TRUSTe, who have been long recognized as the standard bearer of privacy certification, have begun work on a short-form privacy summary [88], though as of this writing the

work seems to have been postponed. TRUSTe does offer a mobile privacy policy short display, similar to offerings from other market players such as PrivacyChoice [90].

Other projects have similarly attempted to create standardized short labels. A Kickstarter project in 2012 launched by Joe Andrieu, collected only half of its $12,000 goal to create a "Standard Information Sharing Label" [3]. None of these formats have yet seen widespread adoption.

2.3.5.1 Privacy Icons

An oft proposed solution to the "problem" of online privacy policies is to replace these policies with a series of icons [92]. We believe these icons to be largely misguided for three reasons:

1. The large number of privacy terms and concepts expressed in privacy policies.
2. The difficulty of expressing these often abstract concepts in simple icons.
3. The end need to expect users to recognize the icons, link them to the concepts, and make comparisons.

As far as we are aware, the first complete set of privacy icons was proposed by Mary Rundle at the IGF Athens Privacy Workshop in 2006 [98]. Her seven icons, displayed in Figure 2.5 were not meant to replace privacy policies but as a hook to display common practices.

Efforts from there ballooned, with designers trying to capture an ever larger number of terms. Matthias Mehldau released a set of icons in 2007, modeled visually directly after Creative Commons, with 30 icons [79]. In 2009, graduate students at Berkeley created KnowPrivacy, to grade and code privacy policies, and came up with a 13 icon scheme to explain their grading [47].

In 2010, Aza Raskin, then at Mozilla, attempted to capture broad support to standardize on a privacy policy format that would be machine readable and displayed through icons [93]. His icons gained the most widespread attention, though suffered from all of the issues above. Raskin left Mozilla, yet they have continued their icon work with two more versions. The first was designed with Disconnect and has a limited number of concepts analyzed [24]. The second group does not have any graphic elements, instead consumers must associate letters with concepts, like "N" for non-personal information and "G" for (geo)location [73]. The

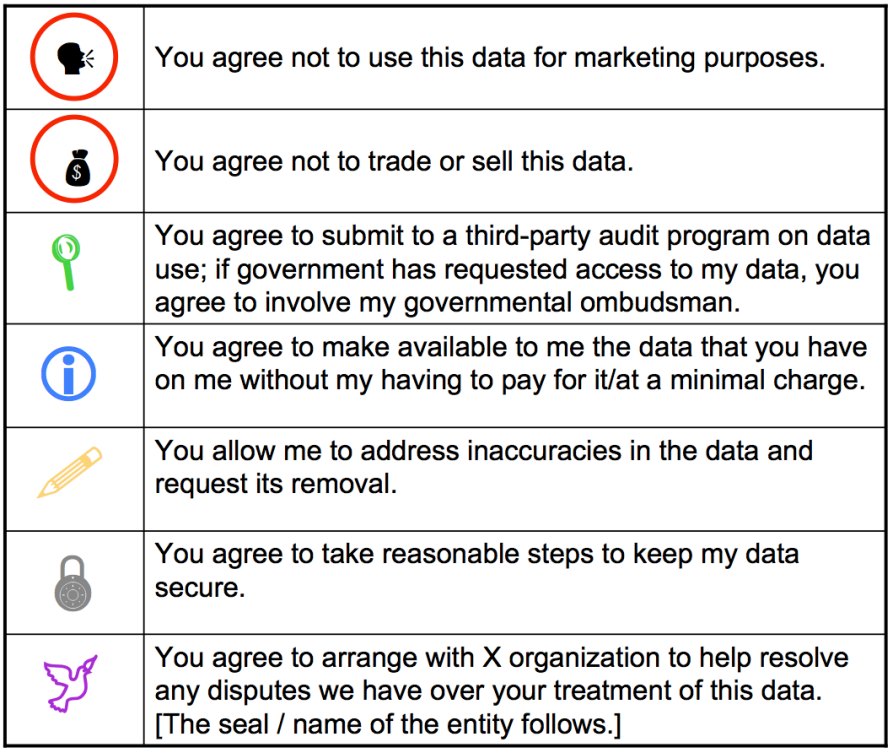


Figure 2.5: An early example of privacy icons, proposed by Mary Rundle, with seven distinct icons, colors, and associated descriptions.

consumer understanding (and misunderstanding) of behavioral icons has also been studied [69].

All of these early projects are currently inactive or operating at very low activity (the Mozilla work has not been updated since September 2011), though new projects do continue to arise [52, 113]. While we continue to see activist demand for privacy icons continue, we believe that these projects will continue to fail until the issues above are reasonably addressed.

2.3.5.2 Terms of Service and End User License Agreements

A number of interesting proposals and research activities are exploring the display of terms and conditions. While not necessarily privacy related, many of the same user comprehension hurdles appear in this domain as well: long text documents, difficult and unfamiliar terminology, and a goal of legal protection. Additionally, these documents are often bound with privacy policies (both online and in users' minds).

Kay's work explored improving these software agreements in two separate ways, the first looks at narrative pictograms, and the second, redesigns agreements to be "textured," highlighting more

relevant information while leaving the text intact [56]. Kay goes on to claim that "Textured agreements achieve improved reading times and comprehension by operating on the primary object of interest—the agreement—rather than creating secondary objects, such as summaries." While this does leave the free-form text in place, it does not address the likelihood that many people do not read these terms at all.

A more recent site, "Terms of Service; Didn't Read" (ToS;DR) admits that these terms (and privacy policies) are not being read and attempts to grade sites into separate classes, A to E, through a "transparent and peer-reviewed process" [97]. While they have yet to rate a large number of sites, they do provide a set of semi-standardized short phrases describing the terms and policies, providing users with a short form label to read. This information would still need to be conveyed to users at a time when they can make a relevant decision. To do this, ToS;DR has created a browser plugin which displays an updating icon to inform the user of the quality of the Terms of Service of every site they visit. Given the green/red system they are currently using, the browser plugin is necessary as websites with low ratings would be unlikely to host those ratings themselves.

Finally, CommonTerms looks to solve this problem at decision time, with an iconified, standardized view of terms and privacy policies, based on a large review of current policies and other iconography and standardization work [68]. While their design and implementation remains in early stages as of this writing, a hybrid approach between an organized set of standardize terms like CommonTerms and a crowd-sourced, yet verified system that allows simple lookups like ToS;DR would likely have the greatest impact.

A "NUTRITION LABEL" FOR PRIVACY POLICIES

EXPERIMENT

3

We used an iterative design process to develop a "Nutrition Label" format for online privacy policies. Drawing from nutrition, warning, and energy labeling, as well as banking privacy notifications, we present our process for constructing and refining a label tuned to privacy.[1]

3.1 RESEARCH QUESTIONS

1. Given the ways today's privacy policies fail (reading level, weasel words [2], term comprehension, length), how can we design a format which remedies these issues?

2. What modifications or simplifications to an industry approved standard (P3P) must we make to provide a foundation to our design?

3. Through our iterative approach we tested a series of possible design choices. What features of our formats, such as color, iconography, length, and terms do users report to be most beneficial?

2. "Weasel words are those whose meaning is twisted, usually for self-serving purposes, like 'popularly priced' – popular to the seller probably." – Gerry Rising, The Buffalo News, March 15, 1999

[1] Portions of this chapter first appeared as "A 'Nutrition Label' for Privacy" [57].

3.2 INTRODUCTION

Website privacy policies are intended to assist consumers. By notifying them of what information will be collected, how it will be used, and with whom it will be shared, consumers are, in theory, able to make informed decisions. These policies are also meant to inform consumers of the choices they have in managing their information: whether use of their information or sharing with third parties can be limited, and if it is possible to request modification or removal of their information.

Today's online privacy policies are failing consumers because finding information in them is difficult, consumers do not understand that there are differences between policies, and policies take too long to read. We set out to design a simplified, standardized, short summary of a company's privacy policy that would help to remedy each of these three concerns.

Our approach comes from a broad survey of work (Chapter 2) that provides consumers with information: nutrition labeling, drug facts, energy information, and most recently work commissioned by the Federal Trade Commission to create a standard financial privacy notice. We discuss our iterative design approach, including focus group testing, as we developed and refined our information design over several months. Finally, we describe our 24-participant laboratory study and discuss the results of our initial evaluation.

3.2.1 The Platform for Privacy Preferences

Instead of defining a standardize privacy policy format from scratch, we chose to leverage an already existing specification. Due to the difficulties surrounding the use of text privacy policies, the World Wide Web Consortium created the Platform for Privacy Preferences (P3P) [114]. P3P is a standard machine-readable format for encoding the online privacy policy of a company or organization. Once a P3P policy has been provided, consumers must use a user agent to interpret it into something understandable. Unfortunately, widely available P3P user agents have limited functionality. These include the P3P policy processing elements of common web browsers and a few privacy-specific browser add-ons [21].

To provide consumers with an active tool where they could investigate and explore the privacy policy of a website the CyLab Usable Privacy and Security Lab (CUPS) produced the P3P Expandable Grid. This user agent was based on one of the central tenants of the Expandable Grid design philosophy, aiming to display a holistic policy view [95]. The interface was created to use and

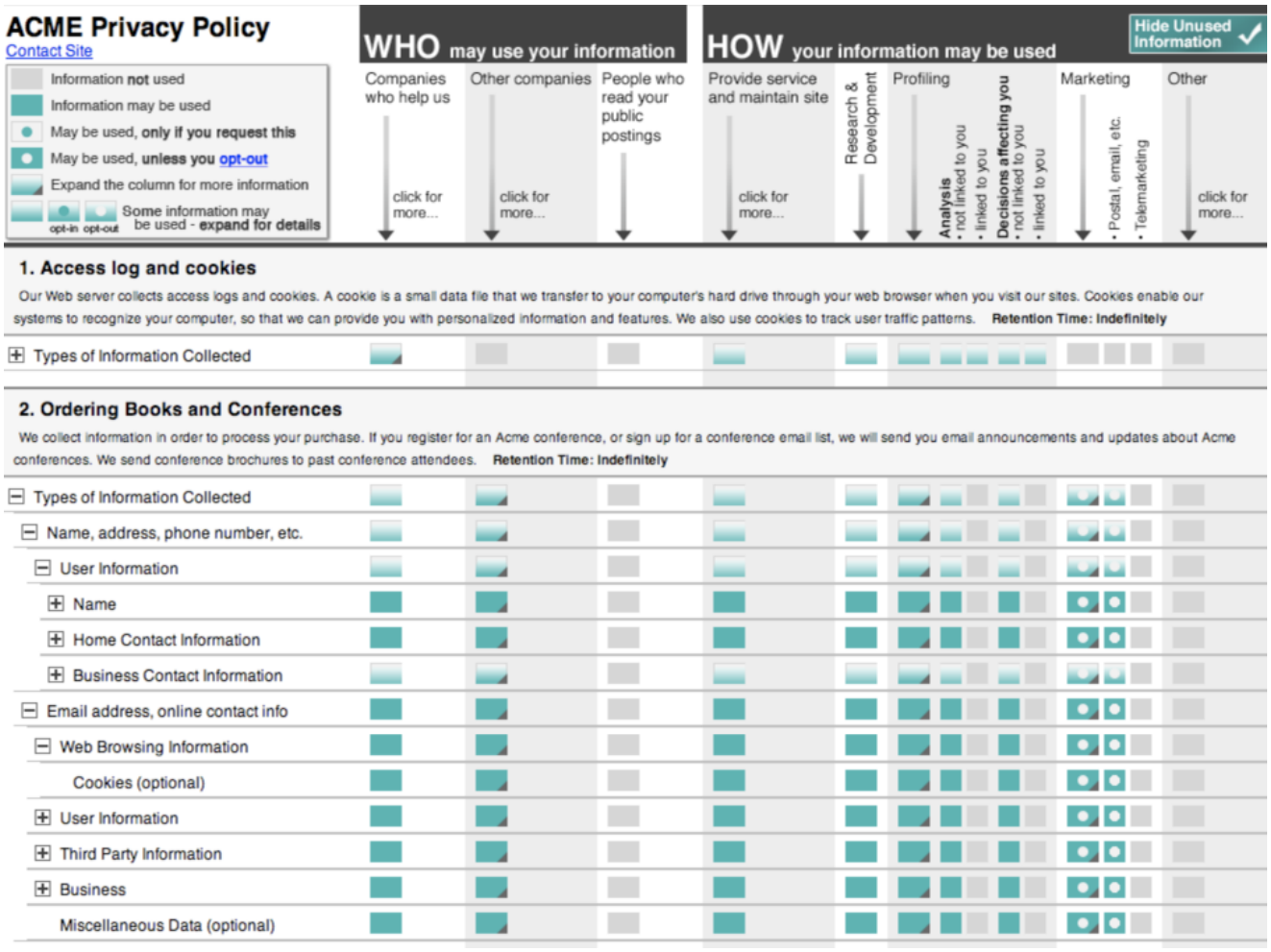


Figure 3.1: Our P3P Expandable Grid, an early attempt at a standardized information design for privacy policies. Due to its implementation of the entire P3P specification, its complexity prevented performance gains for consumers.

display the entire P3P specification, broken down by categories. An example of the grid is shown in Figure 3.1.

The P3P Expandable Grid has two main parts: the header and the body. In the header, there is a title, a legend that explains the 10 possible symbols (8 pictured) that may appear in the body of the grid, as well as expandable column headers that explain how that company uses data and who they will share it with. Finally, in the top-right corner of the header is an additional control that toggles between showing and hiding information that isn't collected (i.e., hide rows that would be blank).

In the body, information is displayed in blocks that correspond to P3P statements. Each block starts with a title and a short textual description (if available) and is followed by a hierarchy of expandable rows, which list what information this company collects. The symbols in each row show how that specific piece of information could be used or shared according to the policy. This format displays

the entire depth of the P3P specification in a two-dimensional grid.

Based on an online survey of over 800 people conducted in the summer of 2007, we found further evidence that people generally do not understand the information presented in privacy policies and also do not enjoy reading them. We compared three formats: a standard natural language policy; PrivacyFinder, which is a simplified human-readable display based on a P3P policy and consisting mostly of bulleted lists; and the above version of the P3P Expandable Grid. None of the three formats were found to be pleasurable to read or easy to comprehend. Notably, we found the P3P Expandable Grid to be slightly worse than the other formats, both in enjoyment and comprehension [94].

3.3 DESIGN METHODOLOGY

This section describes our iterative design process, presenting several prototype labels with benefits and criticisms, and highlighting where knowledge from other label designs was applied. Throughout this process we leveraged informal user feedback as well as focus groups, which are discussed in detail in Section 3.4.

3.3.1 Problems with the P3P Expandable Grid

Based on the analysis of the P3P Expandable Grid study of 520 participants results and a subsequent lab evaluation, we identified five major problems with the Expandable Grid in its current form [60]:

- Many of the standardized P3P labels are not clear to users. For example, “Profiling” and “Miscellaneous Data” are not terms that users encounter in their use of websites.
- The legend has a large number of similar symbols including multiple symbols for expansion (depending on directionality), which the user may not understand.
- Multiple statements that may be related to the same types of information in a P3P policy are displayed separately, possibly requiring the user to check multiple rows to answer a single question. For example, a user might be interested in whether or not a company is collecting their phone number. They may begin expanding rows from the top of the display until they locate phone number, where they see it is not collected or used for any reason. However, they only found the occurance of

phone number in the first P3P statement. To truly know that a company does not collect their phone number they must check each statement. This defies user expectations.

- The Hide Used Information button in the top right only condenses unused rows, not columns.
- Rows with a plus symbol may be expanded; however, almost half of our participants (40.7%) never expanded a single row. By not expanding data types, they never saw some important parts of the policy, necessary for answering the understandability questions we asked [94].

With these initial five problems in mind we abstracted several general principles from the nutrition labeling literature [12, 16, 41, 40].

- All of the labeling efforts we examined use flat, non-interactive designs. While this is partially necessitated by their physicality, an important part of standardizing is deciding how much information should be shown to consumers. Additionally, minimizing the levels and depth of interactivity will simplify the notice for new users.
- Providing a clear and boldfaced title, e.g., Nutrition Facts, communicates the content and purpose of the label and assists in recognition.
- Using bold rules to separate sets of information gives the reader an easy roadmap through the label and clearly designates sections that can be grouped by similarity. Our initial design lacked similar information grouping.
- Putting a box around the label identifies the boundaries of the information. Importantly, this defines the areas that are "regulated" or should be trusted. This is a common issue when the label is placed in close proximity to other information, but may not be as significant an issue online. Our privacy policy formats will all be tested in isolation.

The labeling literature focuses on defining quantifiable properties. This includes the amounts of fats or fiber or percentages of active ingredients or calories based on a standard expected daily value. Privacy policies typically do not include quantifiable measures, and the P3P specification includes no quantifiable fields. The Kleimann Group dealt with this lack of quantifiable information for financial privacy by moving to Yes/No statements, which they found to be

Figure 3.2: Our Simplified Label, an early attempt at a privacy label.

readily understood and reasoned about by focus group participants [64]. While yes and no are not as specific as the number of grams of a substance they do remove considerable corporate wiggle room.

3.3.2 The Simplified Label

Our first design based on other consumer labling efforts was the Simplified Label. In creating the Simplified Label, we used Yes/No statements and applied the four general principles discussed above. An example of the Simplified Label is shown in Figure 3.2. (Note: Each of the screenshots shown throughout this chapter is one of many slight variants of a given format. In each case we have selected one that we believe is representative of the entire series.)

While we made visual changes including adding a title and sub-head, adding bold lines, and simplifying the table view, the most significant change is a reduction in complexity. Two changes contributed most to simplifying the label: eliminating P3P statement groupings and eliminating the use of P3P data hierarchies. These changes are detailed below.

3.3.2.1 P3P Statements

As explained above, P3P specifies groups of information into statements.

This means that all of the collected information in a statement can be used for the same purposes and can be shared in the same way. We can think of P3P as consisting of multiple triplets of information: data, purpose, recipient. We did not include retention because our analysis of over 5,000 unique P3P policies collected by the Privacy Finder search engine [21] showed that the majority of P3P policies state that data is retained indefinitely. In cases where a website has a different data retention policy a note can be inserted at the bottom of the label.

Due to P3P information naturally falling into these triplets, a display such as the list in Figure 3.2 suffers some information loss. For example, it is possible contact information is used for marketing exclusively and purchase information is used for profiling purposes exclusively. Or it is possible that both contact and purchase information could be used for either purpose. By removing the triplets and only displaying a list, we lose that distinction. This will always make privacy policies appear either more or equally permissive than they actually are, never less.

A P3P policy may also have multiple statements. In the P3P Expandable Grid, statements were displayed in a numbered list, separated to show the exact policy in full. In the Simplified Label we merged multiple statements into a single list. For example, consider a policy where the first statement of a policy was about cookies and the second dealt with web activity (including cookies). In the P3P Expandable Grid we would list the categories twice. The first time only cookies would be highlighted; the second, web activity. This would require a user to check two locations to know, for example, all the possible parties with whom cookies would be shared. With the Simplified Label we showed the information from all of the statements in a single list.

3.3.2.2 P3P Data Hierarchies

P3P allows for two interchangeable and different hierarchies of data (specified information for collection).

1. The more commonly used is categories: a list of 17 types of information that companies can collect. When a category is specified a company reserves the right to collect any information that falls under that category (e.g. “Physical

Contact Information" includes name, telephone number, and more).

2. The other data hierarchy, the base data schema, includes every data element that can be specified using P3P, hierarchically arranged (e.g., NAME is a child of USER and includes GIVEN[name], MIDDLE[name], and FAMILY[name]). Further complicating the situation, every element belongs to one or more category (NAME is a member of both demographic data and physical contact information because one's GIVEN name is part of his or her contact information while one's FAMILY name provides demographic information).

In the original P3P Expandable Grid, each category was displayed in its entirety within each statement, with each element of the base data schema hierarchically arranged. This led to nearly 800 elements per category (if fully expanded). To simplify, we display only data categories.

While this afforded us a list of possible information that can fit on a page, it suffered when companies stated they would only collect specific items. For example Contact Information would be displayed similarly if a company collected only one of: a consumer's name, postal address, telephone number, or all of the above information. One way of preserving some of this detail would be to display the specific data elements a company collects when a user clicks on the name of a category, adding a small amount of interactivity for advanced users. Again this design decision leans towards displaying policies as more permissive than they are, never less.

3.3.2.3 Design Notes

To further reduce complexity, information that is not collected or purposes that are not mentioned in a particular policy were not shown. The Show/Hide information button was also removed; thus, there is no way to see uncollected information.

Beginning with this variant, we defined a maximum width of 760px for the label. One important consideration was that the privacy label design be printable to a single page and viewable in the standard width of today's internet browsers. This allows nearly all consumers to see the same label (regardless of browser size) and assists consumers who would prefer to read policies on paper.

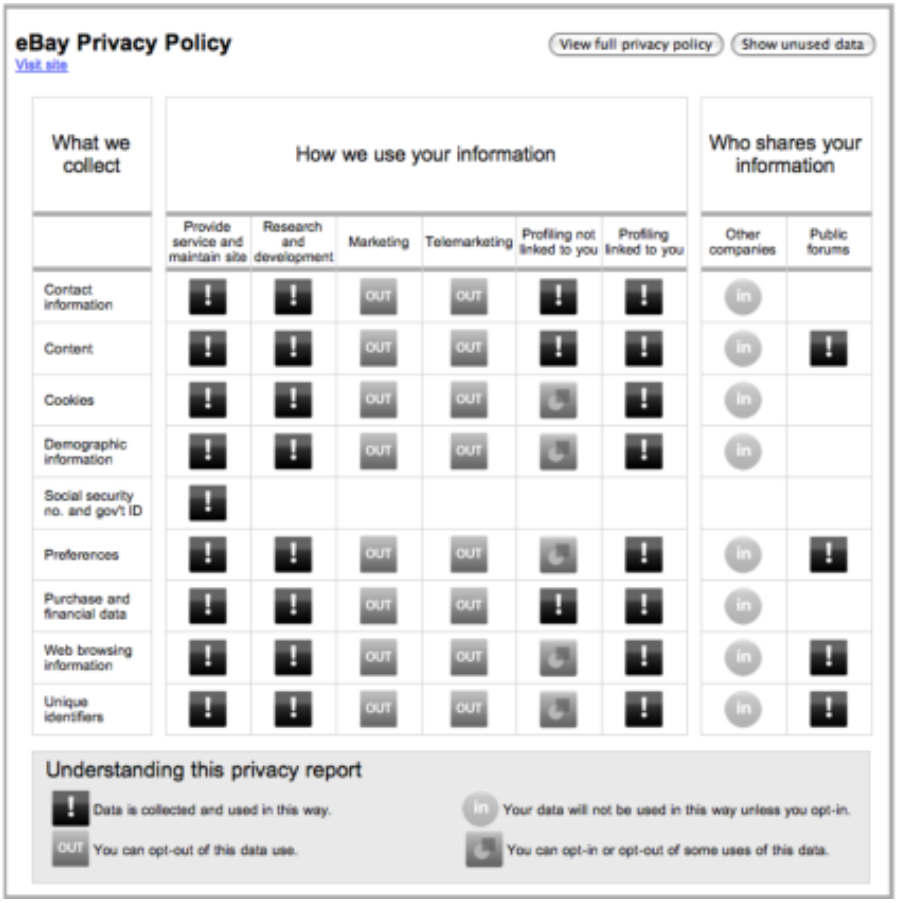


Figure 3.3: Our Simplified Grid in the format returns to a grid base.

3.3.3 The Simplified Grid

While the above label is extremely simple and closely follows a pattern established by the nutrition facts panel and the financial privacy notice, we felt that it sacrificed too much detail.

The goal of our next design was to bring back more of the detailed information that privacy policies can provide without overwhelming users. To do this we tried to find a balance between our Simplified Label and the best aspects of the original P3P Expandable Grid. We adopted a two-dimensional grid layout, as shown in Figure 3.3. We called the resulting design the Simplified Grid.

3.3.3.1 Simplifying the P3P Expandable Grid

While the P3P Expandable Grid was not successful, this failure was not necessarily caused by its table-like display. Also, as discussed above, due to the nature of P3P Statements each reduction in dimensionality causes a loss of information, and we wanted to minimize this loss to create the most benefit for consumers. With

the reintroduction of a two-dimensional layout several changes were made. We continued to only show the data categories, instead of the full base data schema, and also reduced the number of possible recipients and purposes.

There are 12 purposes in the P3P specification, which we merged similarly to the groupings used in the P3P Expandable Grid. Thus, Administration, Current Transaction, and Tailoring are all grouped under the title "Provide service and maintain site." None of these grouped purpose labels are shown. We split the four P3P profiling-related purposes into two categories, based on whether that profiling is linked to the users' identity or performed anonymously. However, during our user testing, this distinction proved unclear to users.

Of the six recipients specified by P3P, Ours and Delivery are both never shown, as it is assumed that the given company will always maintain the information, as well as use it for transaction purposes like delivery. The header "Other Companies" now merges the three remaining types of corporate recipients, distinguished in the specification by their own privacy practices. We decided the importance of this column was to show whether any sharing with other companies was taking place. Sharing with public fora remained unchanged.

3.3.3.2 Symbols and Mixed Control

While you cannot opt-in or opt-out to the amount of trans-fat in a salad dressing you purchase, you might be able to have control over certain aspects of your information sharing on the internet. The Yes/No dichotomy advocated by participants in the Kleimann Group's studies works when there are only one, or maybe two, columns of information. Here we would have needed 8 columns and 10 rows of Yes/No information, which would have been visually difficult to parse.

To solve this, we again looked back to the P3P Expandable Grid and decided to use symbols. However, while the P3P Expandable Grid had 10 symbols, the Simplified Grid uses only four:

-  **Exclamation Point:**
  Data is collected and used in this way.

- 

OUT (in a square):
  You can opt-out of this data use.

-  **in (in a circle):**
  Your data will not be used in this way unless you opt-in.

-  **Square and circle:**
  You can opt-in or opt-out of some uses of this data.

Each of these four symbols was defined in a legend labeled "Understanding this privacy report" directly below the policy. The legend was another device borrowed from the P3P Expandable Grid; however, it was moved below the policy, a less visible place. In this way, users may have to scroll to see it on first use, but returning users will not be forced to scroll past it to see the policy content on future viewings.

Due to the way P3P uses data statements, it is possible that in some instances consumers might be able to opt-out of allowing their demographic information to be used for profiling, but in others it is required, or opt in. The "square and circle" or "mixed choices" symbol attempted to convey this possibility.

3.3.3.3 Visual Intensity

The Simplified Grid was the first iteration of our label to use visual intensity to provide a high-level indication of the quality of a given policy. Each of the four symbols was colored such that darker symbols represented more permissive practices. The use of intensity allows users to make quick visual comparisons that would not have been possible with text alone.

3.3.3.4 Testing

The most significant issue that arose in our testing was confusion over blank areas of the label. We thought that blank areas would clearly indicate information a company does not collect; after all, natural language policies typically leave out any mention of types of information the company does not collect. However, in testing, many participants were unclear on the meaning of the blank cells. Some inferred the accurate meaning that such information uses would not occur, but others thought it allowed the company free rein to do anything in those situations or that they simply had not yet decided their practices.

Additionally, in our user testing the mixed choices symbol was found to be incomprehensible. This served as a further complication on top

Figure 3.4: Our proposed Privacy Nutrition Label. This version was tested in the second focus group and laboratory study.

of the difficulty participants already faced with determining the difference between opt-in and opt-out.

3.3.4 Final Proposed Privacy Nutrition Label

Our Privacy Nutrition Label, shown in Figure 3.5 is a direct descendent of the Simplified Grid. With the Privacy Nutrition Label, we sought to refine the strengths of the Simplified Grid by reducing clutter, introducing color, and simplifying symbols.

3.3.4.1 Types of Information Displayed

We made two changes in the way we present data categories as rows in the table to better facilitate comparisons between policies and to reduce confusion about what data is being collected.

First, all of the P3P Data Categories are now represented in rows regardless of whether they are collected or not. For example, the label shown in Figure 3.5 indicates health and financial information are not collected (and thus not used or shared), but they have not been removed from the display. Second, as a result of this change, every company's policy displayed in this format will have exactly 10 rows, and the ordering will always be consistent. This allows two policies to be easily, visually compared side-by-side.

Participants in a focus group we conducted after making this change did not understand which information companies were not collecting. We indicated the information that was not collected with rows completely filled with minus symbols, but participants believed that companies collected every piece of information listed on the grid. One participant asked, "Why would they collect all that information if they're not going to do anything with it?" In the final prototype we grayed out the labels for data that companies did not collect, and we changed the minus symbol's description from "we will not use your information in this way" to "we will not collect or we will not use your information in this way." We also changed the row-heading label from "What we Collect" to "types of information." This change was made to highlight the fact that we now show even un-collected information and to reduce confusion about what was and was not being collected.

3.3.4.2 Symbol Changes and Color

In the Simplified Grid design, we marked types of information that companies collected and left other cells in the policy blank. However, half of the participants were afraid of the blank spaces; for instance, one said, "Nothing is mentioned. It is completely open-ended. These guys [the company] can modify these values." Therefore, in the final version we introduced a symbol to indicate that information would not be collected or used.

As mentioned, focus group participants found the mixed choices symbol confusing so we removed it. Instead we now display the symbol for the most permissive practice. For example, if in some circumstance one can opt-in and in another one can opt-out, we display the opt-out symbol.

We constrained our initial designs to grayscale to facilitate easy printing without loss of information and to reserve color for highlighting differences between a policy and a user's personal preferences (a natural future extension). However, feedback indicates that color seems to improve user enjoyment in reading the label, although we have not yet quantified this improvement. We selected the colors used in our label with care to accommodate viewers with color-blindness, allow for grayscale reproduction, and maintain the darker- is-more-permissive high-level visual feedback discussed in Section 3.3.3.3.

3.3.5 Useful Terms

Even with the "understanding this privacy policy" legend in place there was still confusion over many of the terms used in the label. This was also an issue during the development of the Kleimann Group's Financial Privacy Notice, and in response they developed what they call the "Secondary Frame." This portion of the prototype notice included both frequently asked questions and a series of extended definitions, which are: "[not] information as essential for consumers to have, but consumers often commented that they liked having it included." [64, p. 27]

Our version of the Secondary Frame is a single page hand-out of useful terms. Our useful terms information was informed by the Human Readable definitions included in the P3P 1.1 Working Group Note [115] and consists of seventeen definitions, one for each of the row and column headers. Some are straightforward, others more detailed. For example, the definition of telemarketing states: "Contacting you by telephone to market services or products," while the profiling definition is:

Collecting information about you in order to:

- Do research and analysis
- Make decisions that directly affect you, such as to display ads based on your activity on the site.

  Information that the site collects about you may be linked to an anonymous ID code, or may be linked to your identity.

In future versions, clicking on or hovering over the headers could pop-up these definitions, but to minimize interactivity and simplify testing these terms are presented below the label and legend.

F A C T S WHAT DOES NEPTUNE BANK DO WITH YOUR PERSONAL INFORMATION?

Sharing practices	
How often does Neptune Bank notify me about their practices?	We must notify you about our sharing practices when you open an account and each year while you are a customer.
How does Neptune Bank protect my personal information?	To protect your personal information from unauthorized access and use, We use security measures that comply with federal law. These measures include computer safeguards and secured files and buildings.
How does Neptune Bank collect my personal information?	We collect your personal information, for example, when you ▪ open an account or deposit money ▪ pay your bills or apply for a loan ▪ use your credit or debit card We also collect your personal information from others, such as credit bureaus, affiliates, or other companies.
Why can't I limit all sharing?	Federal law gives you the right to limit sharing only for ▪ affiliates' everyday business purposes—information about your creditworthiness ▪ affiliates to market to you ▪ nonaffiliates to market to you State laws and individual companies may give you additional rights to limit sharing.
Definitions	
Everyday business purposes	The actions necessary by financial companies to run their business and manage customer accounts, such as ▪ processing transactions, mailing, and auditing services ▪ providing information to credit bureaus ▪ responding to court orders and legal investigations
Affiliates	Companies related by common ownership or control. They can be financial and nonfinancial companies. ▪ *Our affiliates include companies with a Neptune name; financial companies, such as Orion insurance; and nonfinancial companies, such as Saturn Marketing Agency.*
Nonaffiliates	Companies not related by common ownership or control. They can be financial and nonfinancial companies. ▪ *Nonaffiliates we share with can include mortgage companies, insurance companies, direct marketing companies, and nonprofit organizations*
Joint marketing	A formal agreement between nonaffiliated financial companies that together market financial products or services to you. ▪ *Our joint marketing partners include credit card companies.*

Figure 3.5: The Kleimann Communications Group's prototype "Secondary Frame." This page of their prototype notice gives conusmers reasons why companies need to collect and share information (at top) and definitions for common terms (at bottom).

3.4 FOCUS GROUPS

We held two, hour-long focus group sessions to review designs and discuss participants' impressions and questions. We recruited focus group participants from the Carnegie Mellon University (CMU) Center for Behavioral Decision Research (CBDR) participant recruitment website. We paid participants $10 to participate in a 60 minute focus group.

The first focus group was composed of three female and seven male CMU students. The participants reacted positively to the Simplified Grid. For example, one participant stated, "This is more convenient than scrolling through reams and reams of paragraphs. I mean who reads them?" and another participant said, "I like the chart. [It's] better than long sentences." However, we found that some participants still had problems understanding privacy concepts. For example, one participant asked, "What is the difference between opt-in or opt-out?" and many others agreed that they did not understand this distinction.

Many participants had trouble distinguishing different privacy concepts. Most participants were familiar with profiling, but did not understand the difference between "Profiling linked to you" and "Profiling not linked to you." Similarly, participants did not understand the different meanings of "cookies" and "unique identifiers." It was this vein of feedback that led to the inclusion of the useful terms definitions described in Section 3.3.5.

By asking participants to compare two policies, we found that participants could easily isolate and describe differences. Participants noticed that Policy A had more opt-in symbols and Policy B had more opt-out symbols. However, participants were not able to make accurate judgments about the policies. When we asked the participants to chose the company with whom they would prefer to do business, five of the ten participants chose Policy B: the company that collected and used more of their personal information.

Using the feedback from the first focus group, we initiated another series of rapid iteration and prototyping, which resulted in the final label prototype. Our second focus group compared the final Privacy Nutrition Label to the Simplified Label.

The second focus group was composed of four female and three male undergraduate students from CMU and the University of Pittsburgh. We found that participants better understood the grid and were able to make more accurate side-by-side comparisons. Participants understood the significance of the red symbols, saying, "Red is for 'stop' or 'danger'." We passed out two privacy policies, Policy A and

Policy B, and asked the participants to raise their hands if they believed that Policy A is the better policy. Every participant raised his or her hand, correctly identifying Policy A as the more favorable policy. Participants demonstrated a detailed understanding of the differences between the policies with comments such as "It's very clear which site is best" and "You should pick a site with more opt-ins than opt-outs." Some participants even noted subtle differences between the two policies saying, "Policy A isn't perfect either, because they share your preferences, and this may include things like your religious or political preferences."

After reviewing the grid design, we passed out the same policy in the Simplified Label format. Participants reacted negatively to the format because they felt that it did not provide enough information, saying, "This is an empty policy, it says nothing. I wouldn't trust it." Participants wanted to see how each piece of information was being used. For example, one participant stated, "With the grid it's easier to see things. What information is being shared? We don't know that anymore."

3.5 USER STUDY METHODOLOGY

Based on the feedback from our second focus group we performed a 24-participant laboratory user study comparing a standard natural language (NL) privacy policy with privacy policies presented in our Privacy Nutrition Label format.

We used a within-subjects design where participants were randomly assigned to first use either the label or the natural language format. Each participant completed 24 questions relating to the policy format they were shown first and then the same 24 questions again with the other format. These tasks are detailed below. We recorded accuracy as well as time for each participant.

3.5.1 Participants

We again recruited the 24 participants through the CBDR website. Our only requirement was that English be the participant's native language. We offered participants $10 to participate in a 45 minute study in our laboratory. Our participants included 16 students and 8 non-students. Of the 16 students, 5 studied humanities, 5 economics or business, 2 science, and 4 information science. 16 of our participants were male, 8 were female.

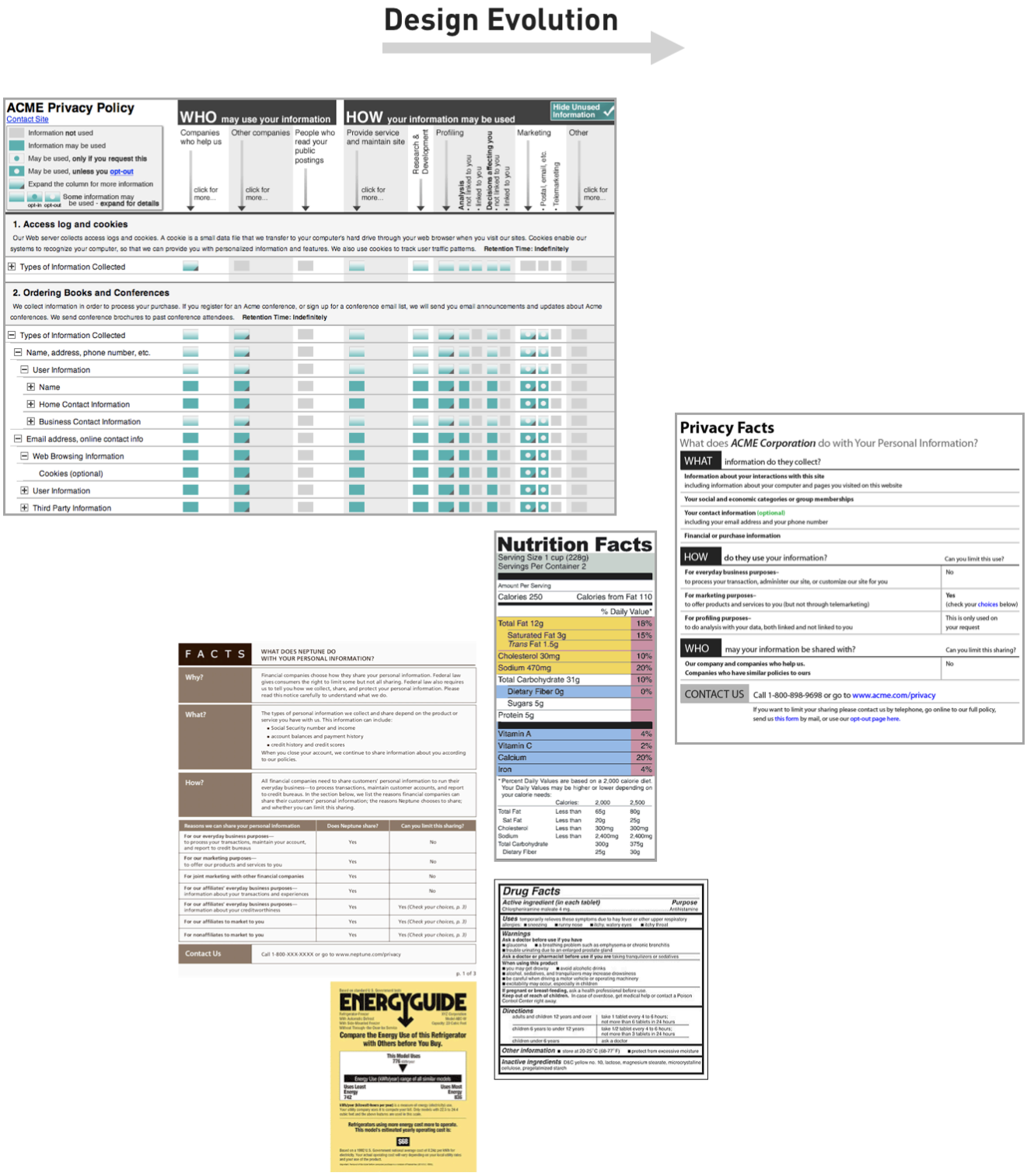
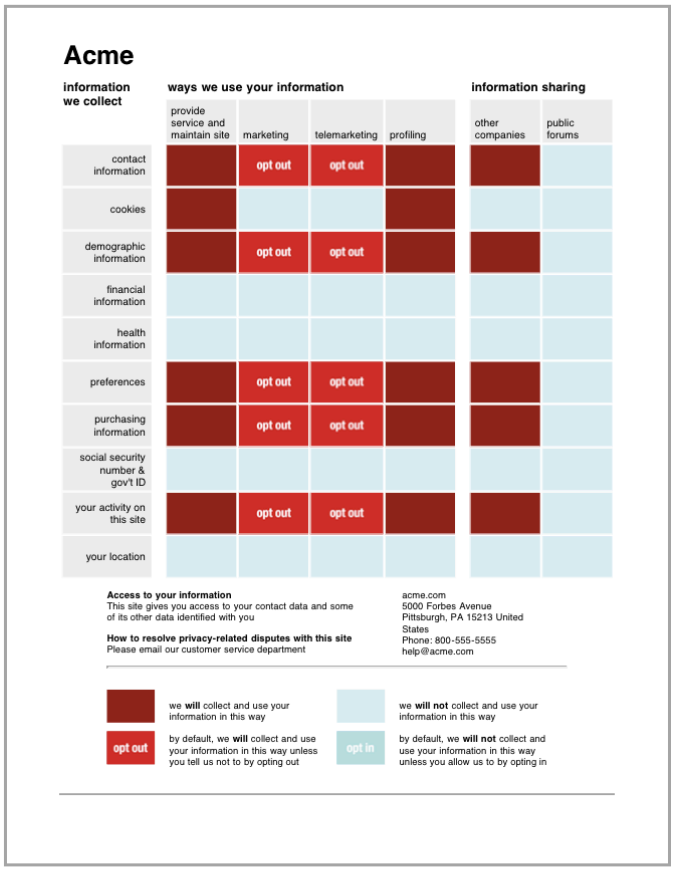
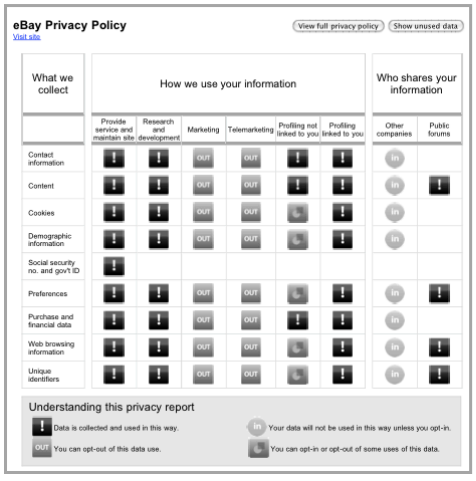


Figure 3.6: Design evolution throughout our iterative process

Final Proposed Design





3.5.2 Privacy Policy Selection

Our study used two NL privacy policies and two label formatted policies. We started with the current actual P3P policy of a popular online e-commerce website. We modified this policy in three ways to produce two slightly different policies for the mythical companies Acme and Button. The first change was to the data collected. Acme has preference information collected but not demographic information, whereas Button Co., collects demographic, not preference. This change does not make one company more clearly invasive than the other. The second change was to the data uses. Acme was modified to not perform any profiling while Button Co. would. The third change was to information sharing practices. While Acme would only share information when consumers opt-in, Button Co. would share information unless consumers opt-out. These differences were introduced so that there would be a clear "correct" response for participant tasks that require them to determine which company better protects their privacy.

The two NL policies for the mythical companies ABC Group and Bell General represent the exact same policies as described above. The ABC Group policy is the natural language policy of the same company whose P3P policy was used to populate the grid, again with the three modifications above made to make it match Acme's. We could not however simply make the three modifications to the policy and also present it as the other natural language option because two different companies, no matter how similar their practices, would not share the same text. The introduction, structure, and actual language used needed to be different. Thus, to create the Bell General policy we used the text of a different, yet comparable e-commerce website, and changed the practices to match that of Button Co.

In editing the natural language policies we removed any references to programs that would distinguish the companies (such as specially branded programs), removed lists of links from the beginning of the policies, removed references to Safe Harbor, and additionally modified the second policy so that both were approximately the same length. For a more complete comparison see Table 3.1.

We chose not to use layered policies. Layered policy adoption is not consistent or widespread, most common layered policies would not be suitable for answering the questions we asked, and recent research has suggested layered policies are no better at helping consumers understand privacy than full natural language policies [78]. We will revisit this decision and perform our own test of layered policies in Chapter 4.

Policy Metric	ABC	Bell
Word Count	2287	2299
Sentence Count	136	130
Flesch Reading Ease	42.06	41.69
Flesch-Kincaid Grade	11.57	11.84

Table 3.1: Extended Text and Readability Comparison for NL

3.5.3 Task Structure

The task structure for each condition was exactly the same, with 24 tasks comprising a section. The 24 tasks were designed in four parts, each of which is detailed here:

3.5.3.1 Information Finding

The first eight questions were all Yes/No questions asked of a single policy (our "A" policies: ABC Group for NL, Acme for the label). Of these 8 questions, 6 were single-element questions, involving only one element of the P3P statement triplet. For example: "Does the policy allow the Acme website to use cookies?" or "Does the policy allow the Acme website to share your information on public bulletin boards?"

The remaining two questions required two elements of the triplet to answer the question, for example "By default, does the policy allow the Acme website to collect your email address and use it for marketing?"

3.5.3.2 Perceived Privacy Policy Understanding

Following the 8 information-finding questions, participants were given 6 questions on a 5-point Likert scale, from Strongly Disagree (1) to Strongly Agree (5). Each of these is described below.

L1: "I feel secure about sharing my personal information with Acme after viewing their privacy practices" attempts to capture participants' reaction to the actual content of the privacy policy they read. L2: "I feel that Acme's privacy practices are explained thoroughly in the privacy policy I read" questions whether participants believe their practices are well displayed.

The next three questions deal with the experience of interacting with the privacy policy in the format we presented. L3: "Finding

information in Acme's privacy policy was a pleasurable experience" has participants rate their enjoyment of using the format. L4: "I feel confident in my understanding of what I read of Acme's privacy policy" investigates participants' perceived accuracy in the information-finding questions. L5: "It was hard to find information in Acme's policy" has participants rate the difficulty they had in finding information.

The final question, L6: "If all privacy policies looked just like this I would be more likely to read them" attempts to capture whether our proposed label would encourage more people to read privacy policies.

3.5.3.3 Policy Comparison Questions

The third section requires participants to compare two policies of the same format (ABC Group v. Bell General for NL or Acme v. Button Co. for the label). One of the policies in each comparison is the same policy from the initial 8 information-finding questions.

The first four questions in this section are True/False statements such as "By default, Button Co. can share information about your purchases with other companies, but Acme cannot."

The final two questions in this section are opinion questions, asking: "Which company will better protect your information online?" and "You're looking to buy a gift online. At which company would you prefer to shop?"

3.5.3.4 Policy Comparison Enjoyment and Ease

The final four questions are again on the 5-Likert scale presented earlier. They are in two pairs, the first pair asking if, "Looking at policies to find information was an enjoyable experience" and "Looking at policies to find information was easy to do." The second pair focuses specifically on the comparison task, "Comparing two policies was an enjoyable experience" and "Comparing two policies was easy to do."

3.6 RESULTS

The results from our laboratory study are presented below. First we will address the issue of information finding through our quantifiable accuracy results. Next we describe the timing data on those questions, showing information finding is not only more accurate but also faster with label polices than with NL policies. To

#	Label	NL	McNemar's p-value	Benjamini-Hochberg Correction
1	96%	100%	NS	NS
2	**88%**	**29%**	**0.00024**	**0.0014**
3	100%	96%	NS	NS
4	92%	100%	NS	NS
5	54%	25%	0.12	0.21
6	**79%**	**21%**	**0.00012**	**0.0014**
7	75%	54%	0.3	0.45
8	88%	58%	0.09	0.18
15	96%	63%	0.06	0.14
16	92%	79%	NS	NS
17	**83%**	**38%**	**0.007**	**0.021**
18	**71%**	**25%**	**0.0009**	**0.0036**

Table 3.2: McNemar's p-values and Benjamini-Hochberg Correction p-values for information finding questions 1-8 (3.5.3.1), and policy comparison questions 15-18 (3.5.3.3).

conclude this section we will present the "likability" of the privacy label.

3.6.1 Accuracy Results

At a high level, people were able to answer more questions correctly with the label. We compared the correct number of total questions, per participant, for the label vs. the natural language policy, $M = 10.13$ and $M = 6.83$ respectively, $t(23) = 7.41$, $p < 0.001$.

We explored each of the questions individually by testing the proportions of correctness for each question by condition, using McNemar's test. These results combine participants who saw the label either first or second as accuracy differences were not significant between these two conditions. These comparisons show that the label is significantly more accurate in 2 of the 8 information-finding questions and 2 of the 4 policy-comparison questions. The accuracy rates for each question are shown in Table 3.2 with statistically significant comparisons shown in bold.

We performed a Benjamini-Hochberg correction to account for multiple testing across comparisons. Each of the paired proportions are shown in Table 3.2 along with the McNemar's p-values and the corrected p-values.

Times in seconds	Label	NL
Information Finding	174.5	349.6
Policy Comparison	120.0	292.4
Average Total Time	339.9	692.0

Table 3.3: Time-to-task comparisons between the label and natural language policies. Shorter times are better. Information finding is questions 1-8 (5.3.1), policy comparison, questions 15-18 (5.3.3)

#	Label	NL	Difference	p-value
1	37.58	61.27	23.69	0.07
2	**21.67**	**85.7**	**64.03**	**0.04**
3	**14.35**	**50.07**	**35.72**	**<.001**
4	18.89	23.09	4.2	0.4
5	34.51	29.95	-4.56	0.46
6	20.19	50.24	30.05	0.06
7	16.32	22.82	6.5	0.88
8	26.93	36.79	9.86	0.73
15	**46.58**	**132.69**	**86.11**	**0.0006**
16	**34.36**	**68.32**	**33.96**	**0.05**
17	21.91	35.48	13.57	0.28
18	**12.24**	**47.36**	**35.12**	**0.03**

Table 3.4: Time differences and p-values for average time per question comparing only correct answers. All times reported in seconds.

3.6.2 Timing Data

For each of the information-finding and policy-comparison questions we collected time-to-task completion data. As shown in Table 3.3, the label was significantly faster than the natural language policies for both the group of information-finding questions and the group of policy-comparison questions ($p < 0.001$).

To test the mean task completion time for accurate answers we removed all timing results where the resulting answer was inaccurate and calculated means per question, per condition. Using a 2-sided t-test the label is significantly faster in 2 of the 8 information-finding questions and significantly faster in 3 of the 4 policy-comparison questions. In only one question was the average time faster for participants using the natural language policy, and this difference was not significant. The full results for this test can be found in Table 3.4.

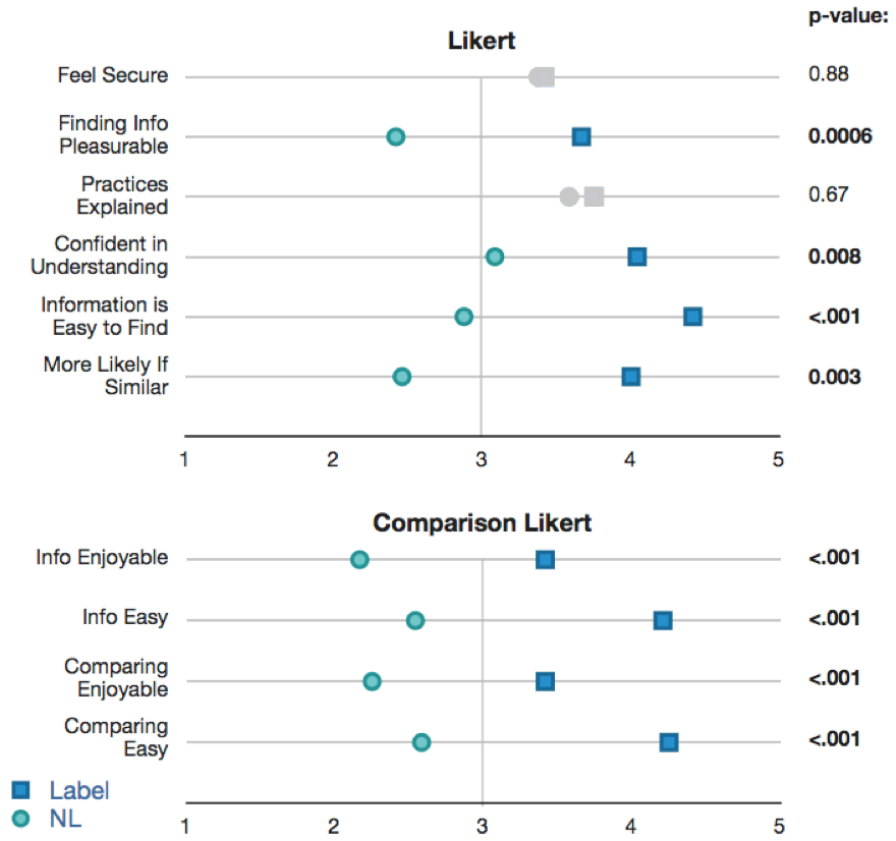


Figure 3.7: Satisfaction results from Likert questions in our laboratory testing. Higher numbers are better, significant results are colored and show bold p-values on right.

3.6.3 Satisfaction Results

The satisfaction results were captured based on participants' responses on a Likert scale from 1 (Strongly Disagree) to 5 (Strongly Agree). We computed the mean response for the label and NL formats, both combined, and also separated by which format was viewed first. For each of these questions higher is better, including Question L5 "information was hard to find," which was reversed to be consistent with the remaining questions.

We performed t-tests for each of these questions, to compare the label to the natural language policies. All but 2 of these 10 questions resulted in significant results. The label was rated significantly more pleasurable, easier to find information in, and easier and more enjoyable to use when comparing two policies. The results from each of these questions are shown with means and p-values in Figure 3.7.

Additionally we performed 2-sample t-tests between conditions to exploring priming effects, where opinions have changed based on the policy format a participant viewed first. When reviewing how participants answered the Likert scale questions about the label by

condition, 3 questions had significant results. Participants felt significantly more secure when viewing the grid if they saw the NL policy first, (label first=2.92, NL first=3.92, $p = 0.03$) reported they were significantly more likely to read policies more in the label format if they saw the NL policy first (label first=4, NL first=4.5, $p = 0.04$), and found comparisons on the label significantly easier when viewing the NL policy first (label first=3.92, NL first=4.58, $p = 0.004$). These results show significant priming to appreciate the grid more when the NL policy was viewed first.

3.6.4 Observations

The initial results from this evaluation were very strong, however there was still much room for improvement. We observed that some participants still found elements of the label confusing. An additional round of iterative design and testing was conducted to address some of the issues we observed during the lab study.

Several participants were confused by the symbols we used to indicate opt-in and opt-out. For instance, one participant did not understand what "out" meant, saying, "I've been messing things up because I thought 'out' meant 'out of the question.'" To improve users' comprehension, we altered the symbol design to include the full phrases "opt-out" and "opt-in."

In addition, several participants in the lab study were completely unfamiliar with the terms opt-out and opt-in, and they assumed that the terms meant exactly the same thing. We updated our glossary definitions to help educate users about these concepts. The original definitions did not explain the terms opt-in and opt-out, with the legend reading "we will collect and use your information in this way unless you opt-out." The refined definitions help explain the concepts, stating: "we will collect and user your information in this way unless you tell us not to by opting out."

3.7 DISCUSSION

We began this chapter with three questions in mind:

1. Given the ways today's privacy policies fail (reading level, weasel words, term comprehension, length), how can we design a format which remedies these issues?

2. What modifications or simplifications to an industry approved

standard (P3P) must we make to provide a foundation to our design?

3. Through our iterative approach we tested a series of possible design choices. What features of our formats, such as color, iconography, length, and terms do users report to be most beneficial?

3.7.1 Can We Remedy The Issues Today's Policies Face?

Today's privacy policies are failing consumers. With an inability to find information, a lack of understanding that there are differences between privacy policies and control, and the simple time-based costs today's privacy policies are not successful. We designed a short summary of a company's privacy policy that helped to remedy each of these three concerns and was at the same time be enjoyable.

We believe that the results presented above clearly show that each of these areas was addressed. Accuracy results were better or similar for information finding, with more participants able to correctly answer questions about the companies' data policies.

Participants recognized that policies were different, were able to make comparisons and answer questions correctly about the differences between example policies, as well as select policies they would prefer to personally use. When using the label people far more consistently selected the company that had the stronger privacy policy. Participants also realized the benefits of the label for comparison: "This may actually be the biggest advantage of this system because you can put down two polices that are formatted the same and see the exact differences between them. It's really easy." Even more directly one participant said "I guess I'll look to see which policy has more blue," leveraging one of our specific design interventions.

Task completion times were significantly lower when using the label than when using a natural language policy. Across the board, participants believed information was easier to find and had a more pleasurable time finding it using the label.

Our prototype label does have limitations. As mentioned the concept of opting in versus opting out is unclear for many of our participants and is the only way our label highlights user choice. Our label summarizes and hides many of the specific details that may be present in a text privacy policy. While users can always read the full text of a company's privacy policy, providing a deeper, more detailed interaction through the label itself should be seen as a topic for future exploration.

3.7.2 How Can We Leverage An Industry Standard?

At first blush, P3P did not seem suited to providing simple, short-form privacy notices. However, by understanding the structure, merging and refining groupings, and using the human-readable definitions, we were able to create a label that can be automatically generated from a P3P file, or filled in by hand by a company. This means that for companies whom already specify P3P policies, a label can be made instantly, no extra work required on the part of the company.

This also means that we are not the arbiters of what types of data companies collect or with whom they share it. By basing our label on an already accepted standard which had legal and corporate involvement, we gain a more solid foundation.

3.7.3 Which design choices will best benefit users?

The final label design allows for information to be found in the same place every time. It removes wiggle room and complicated terminology by using four standard symbols that can be compared easily. While we experimented with more symbols, based on user feedback, and to avoid confusion we reduced this number to four.

It allows for quick high-level visual feedback by looking at the overall intensity of the page. Though our participants did not all recognize this feature, some found it immediately valuable and others were able to leverage the contrast after they were made aware of it. Based on feedback from older consumers and the Kleimann Communication Group's findings the label can be printed to a standard sheet of paper, and also fits in a browser window that will support even older devices.

Even after refinement of our terminology, we found that some users did not have clear definitions for the concepts we were using. As a result we provide a glossary of useful terms and clear definitions for our symbols below the label.

People who have used it to find privacy information rated it as pleasurable. They not only rated it better than the natural language policies that are today's standard, but actually rated it enjoyable to use.

A LARGE-SCALE TEST OF STANDARDIZED PRIVACY LABELS

EXPERIMENT

4

We conducted an online user study of 764 participants to test if more-intentionally designed, standardized privacy policy formats, assisted by consumer education, can benefit consumers. Our results show that standardized privacy policy presentations can have significant positive effects on accuracy and speed of information finding and on reader enjoyment of privacy policies.[1]

4.1 RESEARCH QUESTIONS

1. Does our policy format provide the same improved results that we saw in our laboratory experiment in large-scale testing on:
    - accuracy
    - speed
    - user sentiment
2. Is this format better at facilitating policy comparisons, specifically guiding users to select companies that collect less information?
3. How do layered privacy notices compare to other formats on the above metrics?

[1]Portions of this chapter first appeared as "Standardizing Privacy Notices: An Online Study of the Nutrition Label Approach" [58].

4. Are our formats performing better due to standardization, tabular formats, or both?

4.2 INTRODUCTION

The last two chapters have shown that consumers cannot effectively find information in today's privacy policies and that they do not enjoy using them. In the last chapter we iteratively developed a standardized, table-based format for privacy policies. We compared this format to the common natural-language, full-text policies that exist today in a small laboratory study. With this favorable first comparison we will expand our format here for large-scale verification.

In the experiment presented below, we compared this standardized format, and two variants (one tabular, one text) with the current status quo: full-text natural-language policies and layered policies. We did this through an online user study of 764 participants to test if these the three formats we designed have consumer benefits.

4.3 POLICY FORMATS

We tested five privacy policy formats: standardized table, standardized short table, standardized short text, full policy text, and layered text. The first three of these formats are standardized and were created by our lab using an iterative design approach. Of these, two are tabular and one is textual. Two explicitly describe absent information (below the policy) and one presents it in the context of the policy.

Each of these formats is followed by a list of 16 definitions of privacy terms, consistent across formats. These definitions define the row and column headers in the tables and the text tokens in the standardized short text. They also assist with understanding some of the terminology used in the survey questions.

4.3.1 Standardized Table

The standardized table format, (Figure 4.2, left) has ten rows, each representing a data category the company may collect, four columns detailing the ways that data may be used, and two columns representing ways that data may be shared outside the company. This table is filled using four symbols, dark red to represent that your

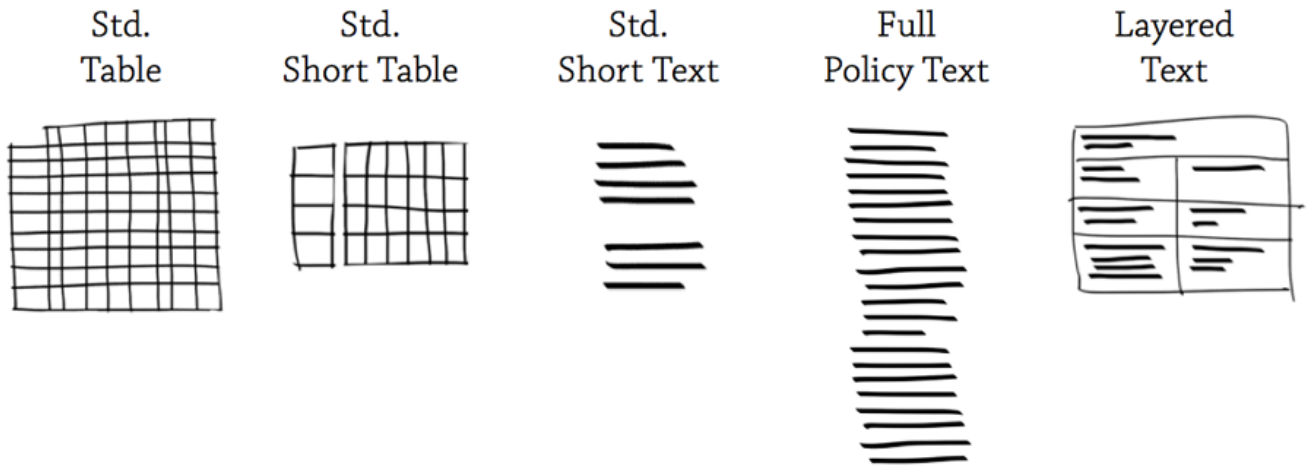


Figure 4.1: Graphical representation of the five different formats.

data may be used or collected in that way, light blue to represent that your data will not be used or collected in that way, and two intermediate options labeled “opt in” and “opt out.” This format should feel familiar, as it is a variant of the “nutrition label” format discussed in the last chapter, modified based on later design iterations.

4.3.2 Standardized Short Table

The standardized short table (Figure 4.2, right) is a shortened version of our proposed tabular approach, which removes the data categories (rows) that are never collected by a company. These removed data categories are listed immediately following the table to maintain a holistic understanding of a company’s privacy practices. By removing the rows, the holistic display requires text comprehension, and cannot solely be performed visually. While the removal of data categories allows the table to fit into a smaller area, it makes comparisons between policies less straightforward due to rows not being vertically matched.

4.3.3 Standardized Short Text

We additionally created a short, natural-language format (Figure 4.3) by translating each row in the standardized short table into English statements, using the column and row headers from the table to form each statement. Rows that have exactly the same sharing pattern across each element are merged into single paragraphs. This means the label will have at least one paragraph, and at most 10.

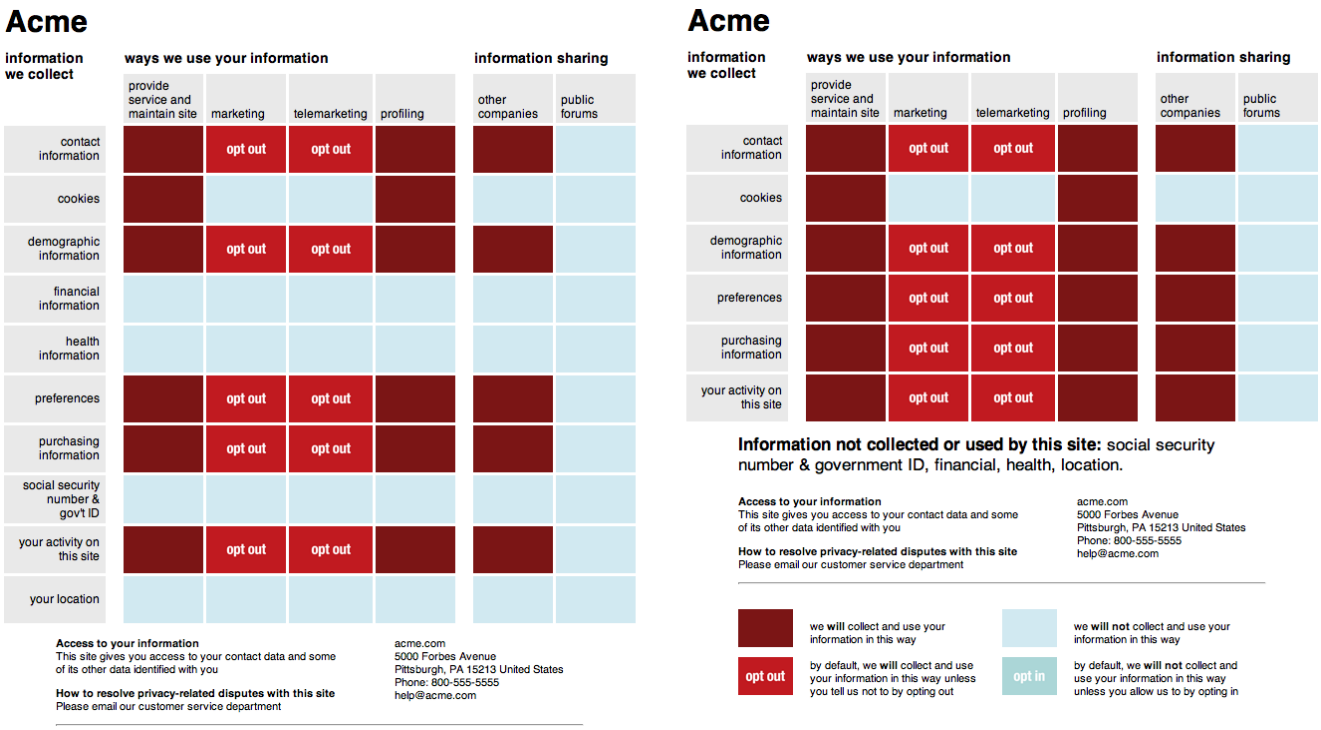


Figure 4.2: An example of a standardized table is shown on the left, and the same policy shown in a standardized short table format on the right. The comparison highlights the rows deleted to "shorten" this version. These deleted rows are listed directly below the table. While both formats contain the legend (bottom right), it is displayed only underneath the standardized short table due to space constraints.

Each paragraph follows a similar pattern. The first sentence specifies the types of information the paragraph is about. The second, third, and fourth, sentences specify what purposes this information must be used for, will be used for unless opting-out, and will be used for only if opting-in, respectively. The final sentence describes the sharing of that information.

Note that this format does not specify data which is collected, and then not used for specific purposes, making it slightly less complete than the two other standardized formats. This format allows us to compare standardized textual and tabular formats directly.

4.3.4 *Full Policy Text*

Natural-language, full-text policies are the de-facto standard for presenting privacy policy information online. For this experiment, we selected four policies from well-known companies. We stripped each policy of all formatting, retaining only internal hyperlinks to reference other areas of the policy, if available in the original. We anonymized all identifying branding, including company and product names, affiliates, and contact information.

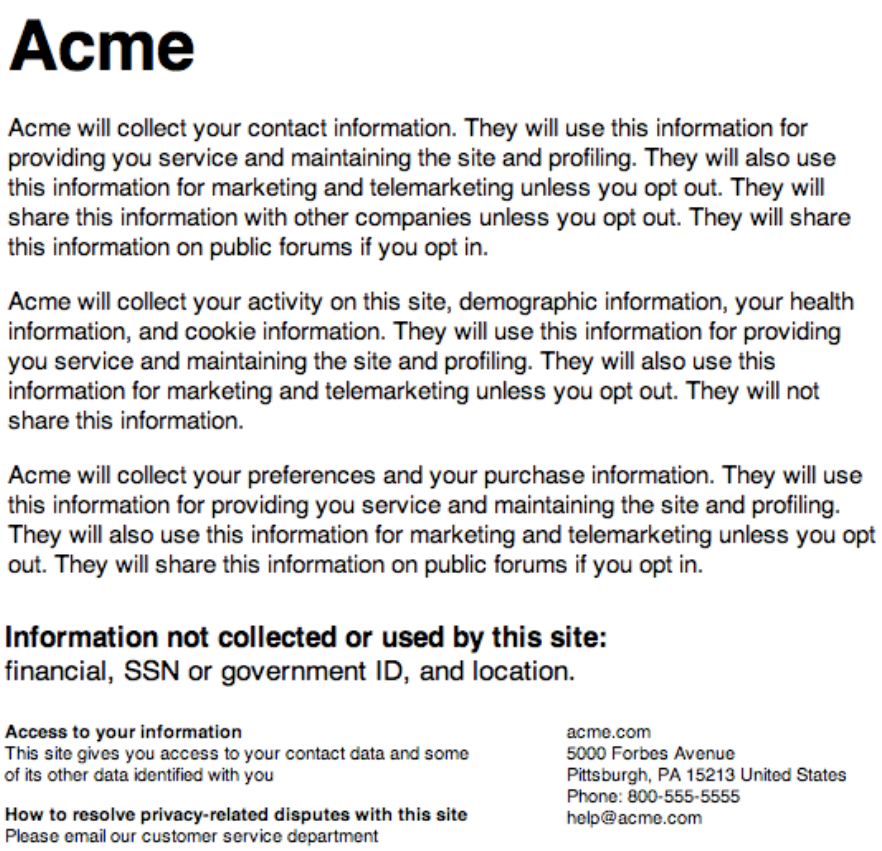


Figure 4.3: An example of the standardized short-text format.

While we have already discussed in depth the problems with full-text policies, we will continue to use them as a baseline in this experiment.

4.3.5 Layered Text

Finally, we tested the layered privacy notice as described by the law firm Hunton & Williams, mentioned earlier — a summarized, one-screen privacy policy in a tabular format that links to the full natural-language policy [104, 105]. Layered policies have been deployed by major corporations, making them a viable, real world summary format for privacy policies. While the use of layered policies has not seen widespread adoption, we wanted to compare them against our metrics to test if they too have benefits against full-text policies. Again we used the policies of actual companies, which were stripped of brand information, but as layered policies are heavily dependent on display styling, we retained the corporate formatting, colors, and text styles.

What Personal Information About Customers Does Amazon.com Gather?

The information we learn from customers helps us personalize and continually improve your shopping experience at Amazon.com. Here are the types of information we gather.

- **Information You Give Us:** We receive and store any information you enter on our Web site or give us in any other way. Click here to see examples of what we collect. You can choose not to provide certain information, but then you might not be able to take advantage of many of our features. We use the information that you provide for such purposes as responding to your requests, customizing future shopping for you, improving our stores, and communicating with you.
- **Automatic Information:** We receive and store certain types of information whenever you interact with us. For example, like many Web sites, we use "cookies," and we obtain certain types of information when your Web browser accesses Amazon.com or advertisements and other content served by or on behalf of Amazon.com on other Web sites. Click here to see examples of the information we receive.
- **E-mail Communications:** To help us make e-mails more useful and interesting, we often receive a confirmation when you open e-mail from Amazon.com if your computer supports such capabilities. We also compare our customer list to lists received from other companies, in an effort to avoid sending unnecessary messages to our customers. If you do not want to receive e-mail or other mail from us, please adjust your Customer Communication Preferences.
- **Information from Other Sources:** We might receive information about you from other sources and add it to our account information. Click here to see examples of the information we receive.

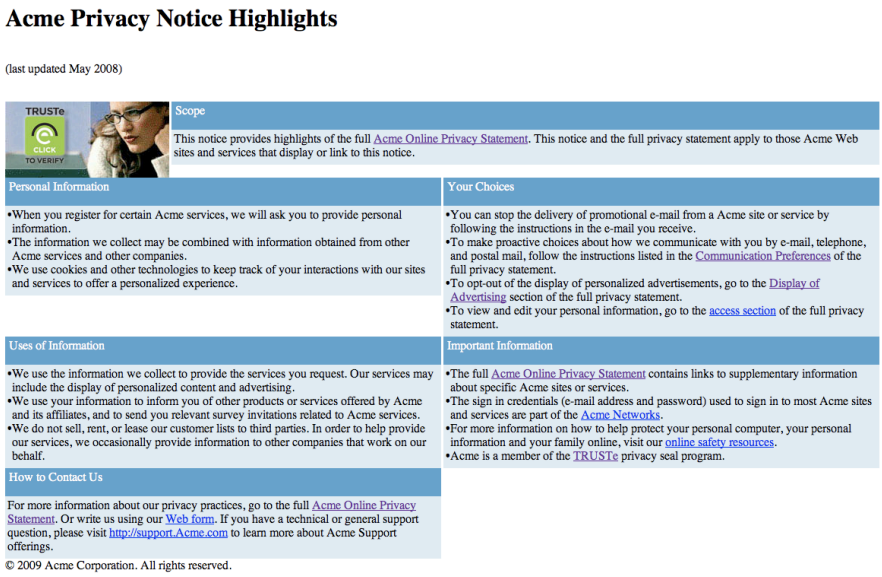
What About Cookies?

Figure 4.4: An example of a full-text policy, shrunk to see length and formatting.

4.4 METHODOLOGY

We conducted an online user study in summer 2009 using Amazon's Mechanical Turk and a tool we developed called Surveyor's Point. Mechanical Turk offers workers the ability to perform short tasks and get compensated. People can place jobs through Mechanical Turk, specifying the number of workers they are looking for, necessary qualifications, the amount they are willing to pay, and details about the task. Mechanical Turk payments are generally calibrated for the length of the task. For our approximately 15-minute study, we paid $0.75 on successful completion.

We developed a custom survey-management tool called Surveyor's Point to facilitate our data collection. Our implementation allows us to show respondents a single question on the screen along with links for switching back and forth between two policies within a single browser window. This allowed us to track the number of users who looked at each policy and the number of times they switched



Acme Privacy Notice Highlights

(last updated May 2008)

Scope

This notice provides highlights of the full Acme Online Privacy Statement. This notice and the full privacy statement apply to those Acme Web sites and services that display or link to this notice.

Personal Information

- When you register for certain Acme services, we will ask you to provide personal information.
- The information we collect may be combined with information obtained from other Acme services and other companies.
- We use cookies and other technologies to keep track of your interactions with our sites and services to offer a personalized experience.

Your Choices

- You can stop the delivery of promotional e-mail from a Acme site or service by following the instructions in the e-mail you receive.
- To make proactive choices about how we communicate with you by e-mail, telephone, and postal mail, follow the instructions listed in the Communication Preferences of the full privacy statement.
- To opt-out of the display of personalized advertisements, go to the Display of Advertising section of the full privacy statement.
- To view and edit your personal information, go to the access section of the full privacy statement.

Uses of Information

- We use the information we collect to provide the services you request. Our services may include the display of personalized content and advertising.
- We use your information to inform you of other products or services offered by Acme and its affiliates, and to send you relevant survey invitations related to Acme services.
- We do not sell, rent, or lease our customer lists to third parties. In order to help provide our services, we occasionally provide information to other companies that work on our behalf.

Important Information

- The full Acme Online Privacy Statement contains links to supplementary information about specific Acme sites or services.
- The sign in credentials (e-mail address and password) used to sign in to most Acme sites and services are part of the Acme Networks.
- For more information on how to help protect your personal computer, your personal information and your family online, visit our online safety resources.
- Acme is a member of the TRUSTe privacy seal program.

How to Contact Us

For more information about our privacy practices, go to the full Acme Online Privacy Statement. Or write us using our Web form. If you have a technical or general support question, please visit http://support.Acme.com to learn more about Acme Support offerings.

Figure 4.5: The layered format is shown, with styles maintained but corporate branding and names removed.

between them. Additionally, Surveyor's Point allowed us to collect the amount of time that users spent reading the policies, as well as information about whether they clicked through to opt-out forms, to additional policy information links, or from a layered notice through to the full-text policy.

In preparation for this study we first performed three smaller pilot tests of our survey framework. We ran our pilot studies on Mechanical Turk with approximately thirty users each, across 2-3 conditions. Our pilot studies helped us to finalize remaining design decisions surrounding the standardized short table, refine our questionnaire, and test the integration of Surveyor's Point with Mechanical Turk.[2]

We then conducted our large-scale study and completed the analysis with 764 participants (409 female, 355 male), randomly assigned to our five policy formats (see Table 4.1): full policy text, standardized table, standardized short table, standardized short text, and layered text. We chose a between-subjects design to remove learning effects and ensure the study could be completed within about 15 minutes. Participants in each condition followed the same protocol; only the policy format differed.

2. The two systems are linked using a shared key that Surveyor's Point generates on the completion of our survey, which a participant then enters back into Mechanical Turk. This allows us to link an entry in Mechanical Turk with an entry in Surveyor's Point and verify the worker completed the survey before payment.

4.4.1 Policies

We selected policies for the study from companies that consumers would conceivably interact with. We narrowed our search by

	Std. Table	Std. Short Table	Std. Short Text	Full Policy Text	Layered Text
Participants	188	167	169	162	78

Table 4.1: Study participants across formats ($n = 764$).

selecting companies that had websites with over one million views per month[3] and were P3P enabled. Additionally, we selected two companies with layered policies deployed on their websites. The four policies we selected were Microsoft, IBM, Target, and Disney, with Microsoft and IBM offering layered policies at the time of the study.

We randomly assigned half our participants in each condition to answer questions about anonymized versions of the Target and Disney privacy policies (Group A), and assigned the other half of our participants to answer questions about anonymized versions of the Microsoft and IBM privacy policies (Group B). By having participants answer questions about policies from different companies we are able to gain insights into where our results may be due to features of a specific policy and where they may be generalizable across many policies.

The policies range in length, but are representative of common practices. Table 4.2 summarizes word counts across the full text, standardized short text, and layered policies.

	Group A Policy 1	Policy 2	*Group B* Policy 3	Policy 4
Full Policy Text	2127	6257	4399	2912
Std. Short Text	175	127	108	90
Layered Text			409	800

Table 4.2: Word counts across the three text variants. Note that the definitions that we append to every policy format add an additional 433 words.

4.4.2 Study Questions

Our study was designed to include questions across seven blocks, with time-to-task-completion recorded for each task:

1. Demographics: We collected standard information about our participants: gender, age, and current occupation.

[3] We used data from `http://www.quantcast.com/` to select these websites.

2. Internet and Privacy: We asked participants four questions to better understand their internet usage and their prior knowledge of privacy.

3. Simple Tasks: We showed participants the "Acme" policy and asked six questions pertaining to it. We refer to these information-finding tasks as simple questions as each question can be answered by looking at a specific row or column in the table. The answer options for these questions (with the exception of question four) were "Yes," "No," or "The policy does not say."

4. Complex Tasks: We asked participants six questions pertaining to the Acme policy. We refer to these information-finding tasks as complex questions because each dealt with some interaction between some category of data and either data use or data sharing. The answer options for these were "Yes," "No," "Yes, unless I tell them not to," "Only if I allow them to," or "The policy does not say."

5. Single Policy Likeability: After completing the simple and complex tasks, we presented a series of 7-point Likert questions for qualitative feedback on the format.

6. Comparison Tasks: We showed participants a notice stating that they would now be comparing two policies: the Acme policy, which they had already seen, with the policy for the Bell Group. We asked three information-finding questions and two preference questions that required looking at both policies.

7. Policy Comparison Likeability: We asked participants three more Likert questions to collect qualitative feedback on the task of comparing two policies.

4.4.3 Analysis

Table 4.3 shows the gender and age breakdown of the participants, as well as the number of privacy policies participants reported reading in the previous six months. 56.4% of our participants reported reading at least one policy in the previous six months. Participants reported that they had the following occupations: student (17.3%); science, engineering, IT (16.5%); unemployed (13.2%); business, management, and finance (9.9%); education (7.3%); administrative support (6.7%); service (4.8%); art, writing, and journalism (4.7%); retired (2.4%); medical (2.0%); skilled labor (1.8%); legal (1.3%); and other (9.3%). 2.7% declined to answer.

	Number	Percentage
Total Participants	764	
Gender		
Male	355	46.5%
Female	409	53.5%
Age		
18-28 years old	321	42.0%
28-40 years old	250	32.7%
40-55 years old	116	15.2%
55-70 years old	31	4.1%
Did not disclose	46	6.0%
Number of Privacy Policies Read in the last 6 months		
Never read a privacy policy	189	24.7%
None in the last six months	130	17.0%
1 policy	100	13.1%
2-5 policies	230	30.1%
5+ policies	101	13.2%
Did not disclose	14	1.8%

Table 4.3: Participant demographics across conditions

While this sample population from Mechanical Turk is not a completely representative sample of American internet users, it is a useful population to study. Our participants appear to read privacy policies more than the general population; however, it is possible that participants, realizing that we were going to ask them to compare privacy policies, may have sought to seem more knowledgeable about privacy policies.

Nutritional and drug labeling literature reports that standardization efforts assist most those who seek out the information [27]. If participants on Mechanical Turk do read more privacy policies than the general population then we may be refining our label to help the group that will be most likely to leverage privacy policy information.

We began our analysis by marking all answers to questions as correct or incorrect (although, as we will discuss later, in some cases there were varying degrees of correctness). We also computed the time it took participants to answer each question. We performed the following statistical analysis:

1. We performed an ANOVA on the average accuracy scores, totaled for each participant, across conditions. We performed additional t-tests for paired comparisons.

2. We also scored each simple, complex, and comparison task individually for accuracy. We performed factorial logistic regressions across the policy formats.

3. We performed ANOVAs on the log-normalized timing information for the above tasks.

4. We performed ANOVAs for the Likert questions.

We excluded participant data from analysis if they did not complete the entire question set. In addition, data from participants who completed the study in less than two standard deviations from the mean of the log-normalized[4] times were excluded. (Group A: $n = 14$, Group B: $n = 11$)

4. Log-normalization is used on analysis of timing information for the remainder of the paper to force a normalized distribution, allowing us to perform ANOVA analysis. Timing information in charts will be displayed in seconds to assist understanding. Mean A: 847 seconds or 14.1 minutes, B: 806 seconds or 13.4 minutes. Cut-off point (2 standard deviations below the mean) A: 268 seconds, B: 262 seconds.

4.5 RESULTS

We begin with a summary of our accuracy results at a high level. The remainder of the section includes a more in-depth analysis of each policy format, a summary of our timing results, and concludes with an analysis of participants' enjoyment of reading privacy policies.

4.5.1 Overall Accuracy Results

Each participant completed 15 information-finding tasks. We scored each participant on a scale from 0 to 15, based on the number of these questions they answered correctly, and averaged those scores across conditions. Note, correct answers varied between conditions since policy content varied between conditions. We present these aggregate results in Figure 4.6. This summary shows a large divide between the standardized and non-standardized formats (ANOVA significant at $p < 0.05$, $F(4, 1094) = 73.75$). The three standardized formats, scoring 62-69%, are shown in light blue; while the two real-world text policies, scoring 43-46%, are shown in red. The standardized policies significantly outperformed the full-text policy (standard table vs. full text, $t(510) = -14.4$, standardized short table vs. full text $t(490) = 12.9$, and standardized short text vs. full text $t(491) = 14.3$, were all significant at $p < 0.05$). The layered format did not perform significantly differently from the full text policy ($p = 0.83$, $t(314) = -0.21$).

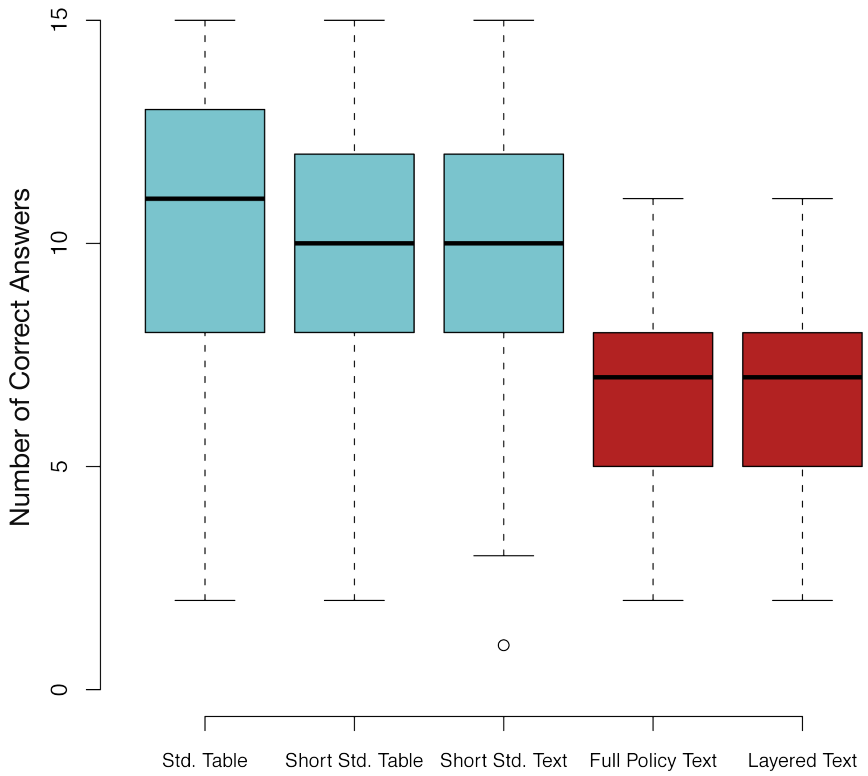


Figure 4.6: Accuracy results for each of the five policy formats.

4.5.2 *Accuracy Results by Format*

We analyzed the results on a per-question basis to gain insights into the strengths and weaknesses of each format. We performed factorial logistic regressions with the standardized table as the base for comparison across formats. The content of the questions and further results are presented in Figure 4.6 at the end of the chapter. Here we describe the overall performance of each format and highlight specific questions to illustrate features of each format.

Where statistically significant results on the user data across formats are discussed in the section below, the values of the statistical tests ranged from $z = 1.97$ to $z = 7.52$, and the level of significance was at least $p < 0.05$.

4.5.2.1 *Standardized Table*

All of our standardized formats benefit from structured information presentation, clear labeling of information that is not used or collected, standardized terminology to minimize length and increase the clarity of the text, and definitions of standardized terms. In addition, the standardized table's tabular display presents a holistic

view of the policy.

Overall, we did not find significant differences between the standardized formats. On a per question basis, the standardized table significantly outperformed at least one of the other standardized formats for some policies in 10 of our 13 questions (3, 5-8, 10, 12, 14-16). It was only significantly outperformed twice (9, 16).

There were some questions that proved difficult across all of the standardized formats. For example, question 3 asked: "Does the policy allow Acme to collect information about your current location?" This information is not collected in any of the policies. The standardized table displays a blank row explicitly labeled "your location," yet more than half our participants answered incorrectly. We believe this was due to participants' misunderstanding the difference between a permanent address ("contact information") and current location information.

Participants performed much better on the other five simple tasks, achieving accuracy results from 73% to 90% across those questions. For the complex questions we saw accuracy results drop; however, the standardized table still fared much better than the full policy text. Question 15 asked: "Does either company collect sensitive information (such as banking or medical records)?" This question was quite complex, involving multiple data types and comparing two different policies. We found overall accuracy between 72% and 84% for people who had already correctly answered question 6 (concerning only medical information). Participants in the standardized-table condition responded correctly three times more frequently than those in the full-policy-text condition.

4.5.2.2 Standardized Short Table

Unlike the standardized table, the standardized short table does not show blank rows for uncollected data categories. Instead the standardized short table lists these categories in a textual notice below the table. The standardized short table showed highly similar overall results to the standardized table. It did perform significantly better for one question (16) and significantly worse for four (3, 6, 14, 15). Importantly, this format still performed significantly better overall, than the full policy text.

This table variant was created to reduce the size of the table, under the assumption that a less space-consuming table could be an advantage over the large table in some applications. However, we were concerned that removing rows would make policies more difficult to compare. The full table performs significantly better in

two (14, 15) of our three comparison questions; the standardized short table only in one (16).

The text notice underneath the table that describes the absence of information requires further isolated testing. Question 6, which asks about the collection of medical information, saw the standardized short-table format perform poorly (59%) when medical information was absent. However, the standardized short-text format, which used an identical notice performed best (81%). Since both formats presented missing information in an identical manner using the same font size, we expect that the difference is due to users being less likely to notice the text when it is under the table than when it is presented in a font size larger than the main policy text on the page.

4.5.2.3 Standardized Short Text

The standardized short text is a direct translation of the standardized short table into text. Rows are grouped by similarity and transformed into sentences. This format did not perform significantly differently from the standardized table overall, but was significantly outperformed in eight questions (5-8, 10, 12, 14, 16), and performed better than the standardized table in only two questions (9, 16). Similar to the other standardized formats, it performed significantly better than the full policy text overall. The standardized short text format is the simplest format we tested. It is compact and requires a participant to understand no symbols, colors, or tables.

One drawback of the standardized short text format is that the length of the text grows with the complexity of the policy. In the longest of our standardized short text policies, the text "Cookie information" is in the middle of a substantial block of text. Only 73% of participants assigned to this policy answered question 5 ("may Acme store cookies on your computer?") correctly (compared to 80-96% across other standardized conditions). At only 175 words, this text seems quite short and participants may not use the search functionality of their browser to find the word "cookie." This is speculative; however, and a follow up study with the paragraphs rearranged may lead us to better understand whether blind spots exist in this format.

We saw evidence of participants in the standardized short text condition misreading text in other questions as well. Question 8 asked: "Does the policy allow Acme to share your home phone number with other companies?" Looking at the standardized short text condition responses for a policy where the answer was "Yes," we see that only 20% answered correctly, while 47% answered "Yes,

unless I tell them not to," implying they believed an option existed where it did not. We suspect this comes from misreading the text, as an option was mentioned for another type of data later in that same paragraph. For question 14 we see a similar pattern. The standardized short text received only 20% accuracy, with 49% of respondents incorrectly answering that neither company gave options regarding cookies.

While the standardized short-text format does perform well, drilling down into these questions shows that it may not scale well, with complex options resulting in longer paragraphs with confusing details.

4.5.2.4 Full Policy Text

As discussed above, the full-policy-text format had worse overall accuracy scores than the standardized formats. We have seen that this format, the de-facto standard on the internet today, has led participants in our study to search multiple pages of text for the absence of a single point of information; uses terms, descriptions, and definitions that may be hard to find or confusing to consumers; and leads to searching multiple sections to find answers which in other formats prove to be much more attainable.

Question 4 asked: "Based on the policy will Acme register their secure certificate with VeriSign or some other company?" Participants in the three standardized conditions correctly answered, "The policy does not say," with 79-88% accuracy. However accuracy dropped to 31-52% for the full-policy-text and layered-text conditions. Neither policy mentioned VeriSign nor any other certificate registrar, nor did either policy contain the word "certificate." We attributed this difficulty to the full-policy-text format forcing users to scan for the absence of information over several pages of text. Today's real world formats simply prevent knowing what will not be done with information.

Worse yet, finding information by looking for specific terms proved difficult. Question 6 asked if medical information was collected, yet in one policy only 49% of participants correctly identified that medical information was collected even though the policy referenced "counseling from pharmacists," an "online prescription refill service," and "prescription medications."

Moving on to complex tasks, question 7 asked: "Does the policy allow Acme to share some of your information on public bulletin boards?" In a tabular format, this question required the participants to find the column for public sharing, and see if any type of data would be allowed. Across the standardized formats accuracy ranged

from 59% to 76%. Participants given the full-policy-text format have strikingly low results for this question (16-34%), regardless of the policy they were assigned. Many incorrectly reported the policy did not specify whether the information would be shared on public bulletin boards, indicating they were unable to find the section of the policy that discussed this.

The first content comparison task, Question 14 asked: "Does either company give you options with regards to cookies?" For the full policy text, 55% of the participants reviewing one set of policies believed that both companies provided options regarding cookies. This means that they incorrectly answered that the Acme policy had options regarding cookies (it did not). Searching for "cookie" in that text brings up a section entitled "Use of Cookies," under which the fourth paragraph reads: "You have the ability to accept or decline cookies. Most Web browsers automatically accept cookies, but you can usually modify your browser setting to decline cookies if you prefer…" Although this sounds like an option regarding the use of cookies, it is not one that Acme provides, rather it is a function of most web browsers. The next line of the policy states "If you choose to decline cookies, you may not be able to sign in…," making it obvious that Acme sites will use, and do require, cookie information. The text in this case, as in many other instances, was confusing.

4.5.2.5 Layered Text

Layered policies, by design, do not provide a complete understanding of a company's practices. Each company decides what information is most relevant to include. Furthermore, companies may use the same language that exists in the full policy text — language that was problematic in the full policy text condition. The layered-text format did not strongly differentiate itself from the full text policy in any of the detailed per-question results that we examined. Given that the participants in this condition had access to the full-text policy we expected this, though only 25 of 78 participants ever clicked through to the main text. While the idea of a short, one page summary, was one of our own design goals it seems this format should be beneficial. However, to create a format that makes a companies' practices understandable, more guidelines around required information seem necessary.

4.5.3 Timing Results

We examined completion times for the simple, complex, and comparison tasks, as presented in Table 4.4. Note, time for

Average Timing Information (in seconds)

#	Std. Table		Std. Short Table		Std. Short Text		Full Policy Text		Layered Text		F-Statistic (dof)
	avg.	σ	avg.	σ	avg.	σ	avg.	σ	avg.	σ	
1-6	236	205	210	103	237	174	367	248	317	406	15.994 (4,756)
7-12	176	194	135	73	163	122	249	358	186	210	8.751 (4,756)
13-17	158	125	148	97	169	122	236	227	187	157	5.094 (4,756)
Full Study	912	572	852	407	938	515	1267	810	1089	768	11.273 (4,756)

Table 4.4: Average time per condition in seconds for questions 1-6 (simple), 7-12 (complex), and 13-17 (comparison), as well as total. While there were significant differences across formats, overall significant differences between the standardized formats were not observed.

comparison tasks includes both information-finding tasks and preference questions. We tested statistical significance using ANOVA on the log-normalized time information across policy formats. For each of these three groups of questions, as well as the overall study completion time there were statistically significant differences across policy formats ($p < 0.0001$ for questions 1-6, 7-12, 13-17, and overall). The standardized formats significantly outperformed the full policy text in overall time (standard table vs. full text, $t(348) = 5.36$, standardized short table vs. full text $t(327) = -6.01$, and standardized short text vs. full text $t(329) = -4.55$, were all significant at $p < 0.05$). The layered format was also significantly faster than the full text policy ($p = 0.025$, $t(238) = 2.25$). The standardized formats, on average were between 26-32% faster than the full text policy, and 22% faster than the layered text policy.

4.5.4 Enjoyability Results

For the most qualitative of our measures, we asked the participants how they felt about looking at privacy policies. We asked six 7-point Likert scale questions after they completed the single-policy tasks and three more after they completed the policy-comparison tasks. The results are summarized in Table 4.5. The Likert scale ranged from "Strongly Disagree" (1) to "Strongly Agree" (7), where higher scores indicate more user enjoyment or perceived usefulness of the format. While there were significant differences for nearly all the Likert questions, we will not go into the details of each question, but average across the two groups of questions.

For the single-policy tasks, participants across the board reported that they felt "confident in my understanding of what I read of

Question Number	Std. Table	Std. Short Table	Std. Short Text	Full Policy Text	Layered Text
*1-6**	4.16	4.06	4.13	4.00	4.14
*7-9**	4.84	4.63	4.47	3.83	4.52

Table 4.5: Mean enjoyability scores on 7-point Likert scale for single-policy questions (1-6), and comparison questions (7-9). The Likert scale ranged from "Strongly Disagree" (1) to "Strongly Agree" (7). While participants feel neutral with a single policy, the range widens when comparing policies. Rows marked with an asterisk represent statistically significant enjoyability differences between conditions (1-6: $F(4, 756) = 4.25, p < 0.05$; 7-9: $F(4, 756) = 10.65, p < 0.05$).

Acme's privacy policy." The question with the most significant strength in the single-policy tasks was the final question: "If all policies looked just like this I would be more likely to read them," with the three standardized policies scoring higher than the full policy text.

The three comparison Likert questions show a larger preference for the standardized formats over the full policy text. We asked whether comparing two policies was "an enjoyable experience," was "easy to do," and if participants "would be more likely to compare privacy policies" if they were presented in the format they were assigned. The gap between the full policy text and the standardized formats widens from about half a point when looking at a single policy to as much as one and a quarter points after making comparisons.

While the layered text notice performed quite similarly to the full policy text in accuracy measures, we see a very different result in participants' feelings about using layered notices. The likert scores for layered policies were not significantly different than the standardized-table format (1-6: $t(756) = -1.57, p = 0.115$; and 7-9 $t(756) = -1.48, p = 0.138$).

The comments provided by participants at the end of the study provide insights into their enjoyment. Participants who saw the full policy text described privacy policies as "torture to read and understand" and likened them to "Japanese Stereo Instructions." On the other hand, participants in the standardized-format conditions were more complimentary: "This layout for privacy policies is MUCH more consumer friendly. I hope this becomes the industry standard."

4.6 DISCUSSION

We began this section with a series of research questions:

1. Does our policy format provide the same improved results that we saw in our laboratory experiment in large-scale testing on:
    - accuracy
    - speed
    - user sentiment
2. Is this format better at facilitating policy comparisons, specifically guiding users to select companies that collect less information?
3. How do layered privacy notices compare to other formats on the above metrics?
4. Are our formats performing better due to standardization, tabular formats, or both?

4.6.1 Do standardized formats hold up in large-scale testing?

Our large-scale online study showed that policy formats do have significant impact on users' ability to quickly and accurately find information and on users' attitudes regarding the experience of using privacy policies.

The three standardized formats that we designed with usability in mind performed significantly better across a variety of measures than the full-text and layered-text policies that currently exist online today. The large amount of text in full-text policies and the necessity to drill down through a layered policy to the full policy to understand specific practices lengthens the amount of time and effort required to understand a policy. Additionally, more complex questions about data practices frequently require reading multiple sections of these text policies and understanding the way different clauses interact, which is not an easy task.

The standardized formats performed the best overall, across the variety of the metrics we looked at. The accuracy, comparison, and speed results eclipse the results of the text formats in use today.

4.6.2 Can this design make policy comparison more accessible?

The standardized formats do help users more accurately answer policy comparison questions. Though given that the policy comparison questions were at the end of the experimental tasks, many participants in the full-text condition seem to have become

burnt out, quickly answering the comparison questions, at a rate that seems highly unrealistic given the length of the documents they were assigned. The much shorter length of the standardized policies may have been an even greater benefit in this task.

The standardized table and standardized short table overall performed very similarly. While there are five cases where the full table outperforms the short table, and only one in the other direction, these differences are frequently small. One concern in the design stage was that removing rows from the table would make comparisons a more cognitively difficult task. This may be evidenced from the significant performance differences in questions 14 and 15; however, the differences in number of rows in the policies we selected were not extreme, never differing by more than one row. It is not clear how great the differences in the types of data collected between real-world policies actually are.

Additionally, we did not test multiple policy comparison, of say a consumer deciding between three, four, or more companies. In theory, the standardized format should scale well with such a task, where we have already seen user frustration with full-text policies when given only two to compare. We will leave this for future work.

4.6.3 How do layered formats compare?

One area where the full-text and layered policies did perform as well as the other formats was on user enjoyment of the single-policy tasks. This may be partially attributed to users' pre-existing familiarity with similar formats. However, this dropped when users reached the comparison tasks, which we expected to be difficult with long text policies. From our earlier work, we observed that when asked to compare the enjoyment of reading policies between the standardized-table format and the full policy text, we noted steep improvements in enjoyment of the table format. With this study's between-subjects design, we were not able to measure such effects, although the free response comments provide some evidence.

Enjoyability results for the layered policies were significantly better than for the full-text policies, even though accuracy scores were not significantly different. This may provide the rational between some corporate adoption: happier consumers who understand just as little about a companies' practices.

Layered policies also took participants less time to use, on average, than full-text policies, although they still took significantly longer than the standardized formats. Some questions could not be answered correctly from reviewing the layered policy without

clicking through to the full policy. However, in this study only 25 of the 79 layered-format condition participants ever clicked through the layered policy to access the full policy. Those who accessed the full policy at least once took an average of 6.6 minutes longer to answer the study questions than those in the layered-format condition who never accessed the full policy. There were not significant differences in accuracy between layered-format participants who never viewed the full policy and those who did; both groups answered just under half the questions correctly.

4.6.4 Are the benefits coming from standardization, tabular formats, or both?

While Chapter 3 showed that the standardized table performed much better than text policies; it was unclear whether the improvement came from the tabular format or the standardization. We have shown here that it is not solely the table-based format, but holistic standardization that leads to success. Our standardized short-text policy left no room for erroneous, wavering, or unclear text, serving as a concise textual alternative to tabular formats.

Though, while the standardized short text policy we developed was successful for most tasks, it is not as easy to scan as a table. Indeed, one participant in that condition suggested policies could be improved if they were set up "like a chart so you can scan it visually for answers instead of having to take the time to read it."

In addition, the standardized short text format may not scale as gracefully as the tables. The standardized short text did perform significantly worse than the standardized table for some questions. This is evident in the information-collection tasks where users had difficulty finding certain types of information in the short text, especially if it was in the middle of a block of text. Because of the way we generate the text, complex policies are longer than simple policies; however, complexity is often privacy protecting and should not be cognitively penalized. The short text policy could grow to up to ten paragraphs for complex policies, which is a concern for information finding.

We asked users after they completed the survey if they used the "search" or "find" functionality in their web browser and found only 14.6% of the standardized short text format reported they did, compared to 44.5% of full policy text format participants. It is possible that by giving users a seemingly more manageable amount of text, they do not use this feature, which can often assist users in looking for specific information.

The standardized short text policy did perform well with information that was not collected, used, or shared, even in comparison to the standardized short table with which it shares an identical text notice for this information. We believe that the notice about unused information stood out more in the text policy than the short table. In the text policy this notice was larger than the other text. In the short table the colorful table is more likely to attract users' attention than the text below it.

4.7 LIMITATIONS

While this chapter concludes our current research on refined notices for online privacy policies there are still future refinements that can be made to these displays.

We expect with further refinement and more testing around educating users in parallel with the standardized formats we could bring accuracy out of the 62-69% range into a passing grade. Specifically, with training geared towards use of the label we may be able to combat the difficulty users had with complex information-finding tasks. Levy and Hastak support continued efforts into education and context, reporting that "consumers have little prior knowledge and experience with information sharing characteristics of financial institutions, they will find it more difficult to understand privacy notice information unless they are provided with more context than is presented in current notices," [70].

We have left a number of small design issues outstanding. As we continue to run experiments we hope to build some of these questions into tests to be able to understand how large blocks of text, additional notices, and colors for highlighting information play a role in the decision making process.

Future work should also continue to concentrate on not just how to present policy information, but also on how to facilitate comparisons, and multiple comparisons.

Finally, the biggest limitation of this work is that these labels are still, largely, a proof of concept. While we have seen some limited deployment, opening up the specification and creating an editor so that they can be widely deployed and then evaluated in the real world remains the final test of the nutrition label privacy policy format.

User Accuracy Across Policy Formats and Groups (with question information)

		#	Question	Answer	Std. Table	Std. Short Table	Std. Short Text	Full Policy Text	Layered Text
Simple Tasks	*Group A*	1	Does the policy allow Acme to collect information about	Yes	82.35	86.25	91.57	80.23	
	Group B		which pages you visited on this web site?	Yes	87.21	85.06	89.53	92.11	84.62
	Group A	2	Acme might want to use your information to improve their website.	Yes	79.41	77.50	83.13	82.56	
	Group B		Does this policy allow them to use your information to do so?	Yes	76.74	77.01	86.05	**89.47**	64.10
	Group A	3	Does the policy allow Acme to collect information	No	48.04	**23.75**	43.37	**18.60**	
	Group B		about your current location?	No	46.51	**24.14**	53.49	**3.95**	**15.38**
	Group A	4	Based on the policy will Acme register their secure certi cate	The policy does not say	88.23	81.25	84.34	**52.33**	
	Group B		with VeriSign or some other company?	The policy does not say	79.07	82.76	87.21	**43.42**	**30.77**
	Group A	5	Based on the policy may Acme store cookies	Yes	89.22	92.50	**73.49**	91.86	
	Group B		on your computer?	Yes	89.53	80.46	87.21	96.05	88.46
	Group A	6	Does the policy allow Acme to collect information about your medical	Yes	84.31	76.25	**69.88**	**48.84**	
	Group B		conditions, drug prescriptions, or family health history?	No	73.25	**58.62**	81.40	**28.95**	**33.33**
Complex Tasks	*Group A*	7	Does the policy allow Acme to share some of your	Only if I allow them to	75.50	76.25	**59.04**	**15.12**	
	Group B		information on public bulletin boards?	No	61.63	57.47	65.12	**25.00**	**38.46**
	Group A	8	Does the policy allow Acme to share your home phone	Yes, unless I tell them not to	62.75	68.75	67.47	**36.05**	
	Group B		number with other companies?	Yes	68.60	60.92	**20.43**	**14.47**	**14.10**
	Group A	9	Does the policy allow Acme to use your buying history	Yes	53.92	58.75	53.01	62.79	
	Group B		to design custom functionality targeted at you?	Yes	50.00	58.62	**69.77**	64.47	**65.38**
	Group A	10	Does the policy allow Acme to share your cookie	No	69.61	67.50	**50.60**	**16.28**	
	Group B		information with other companies?	No	79.07	71.26	74.42	**26.32**	**44.87**
	Group A	11	Will Acme contact you with advertisements?	Yes, unless I tell them not to	54.90	61.25	55.42	**38.37**	
	Group B			Yes, unless I tell them not to	44.19	49.43	51.16	**14.47**	39.74
	Group A	12	Does Acme give you control regarding their sharing of	Yes	70.59	68.75	73.49	66.28	
	Group B		your personal data?	No	56.98	44.83	**37.21**	**31.58**	**24.36**
Comparison Tasks	*Group A*	14	Does either company give you options with	Only with Acme	58.82	52.50	45.78	**33.72**	
	Group B		regards to cookies?	Only with Bell	63.95	**48.28**	**19.76**	**15.79**	**42.31**
	Group A	15	Does either company collect sensitive information	Acme	64.71	**47.50**	53.01	**20.93**	
	Group B		(such as banking or medical records)?	Neither company	73.26	63.22	80.23	**52.63**	**53.85**
	Group A	16	By default, Acme can collect information about your age and gender	True	59.80	61.25	**34.94**	**19.77**	
	Group B		in order to market to you by email, but the Bell Group cannot.	False, both can	56.98	**74.71**	**77.91**	46.05	**24.36**

Table 4.6: Percentage of participants who answered each question correctly, by policy format and viewed policy group. Group A represents participants who saw Policies 1 and 2, Group B, participants who saw Policies 3 and 4. Percentages in bold indicate statistical differences ($p < 0.05$) for formats compared against the standardized table for that policy. For this analysis two separate logistic regressions were performed, a 1x4 for Group A, and 1x5 for Group B. Differences between companies are not compared. Questions are listed exactly as asked, with the corresponding correct answers for each company.

User Accuracy Across Policy Formats and Groups (with detailed statistics)

		#	Std. Table	Std. Short Table	z-value	p-value	Std. Short Text	z-value	p-value	Full Policy Text	z-value	p-value	Layered Text	z-value	p-value
Simple Tasks	*A*	1	82.35	86.25	0.71	0.477	91.57	1.79	0.074	80.23	-0.37	0.710			
	B		87.21	85.06	-0.41	0.682	89.53	0.48	0.635	92.11	1.01	0.315	84.62	-0.48	0.633
	A	2	79.41	77.50	-0.31	0.755	83.13	0.46	0.521	82.56	0.54	0.585			
	B		76.74	77.01	0.04	0.967	86.05	1.55	0.120	**89.47**	2.09	0.036	64.10	-1.77	0.077
	A	3	48.04	**23.75**	-3.31	< 0.001	43.37	-0.63	0.527	**18.60**	-4.10	< 0.001			
	B		46.51	**24.14**	-3.04	0.002	53.49	0.91	0.361	**3.95**	-4.87	< 0.001	**15.38**	-4.11	< 0.001
	A	4	88.23	81.25	-1.31	0.192	84.34	-0.77	0.441	**52.33**	-5.12	< 0.001			
	B		79.07	82.76	0.62	0.537	87.21	1.41	0.158	**43.42**	-4.53	< 0.001	**30.77**	-5.93	< 0.001
	A	5	89.22	92.50	0.75	0.452	**73.49**	-2.70	0.007	91.86	0.612	0.540			
	B		89.53	80.46	-1.65	0.100	87.21	-0.48	0.635	96.05	1.52	0.128	88.46	-0.22	0.826
	A	6	84.31	76.25	-1.36	0.173	**69.88**	-2.32	0.020	**48.84**	-4.98	< 0.001			
	B		73.25	**58.62**	-2.02	0.044	81.40	1.27	0.204	**28.95**	-5.43	< 0.001	**33.33**	-4.97	< 0.001
Complex Tasks	*A*	7	75.50	76.25	0.119	0.905	**59.04**	-2.37	0.018	**15.12**	-7.52	< 0.001			
	B		61.63	57.47	-0.56	0.578	65.12	0.47	0.635	**25.00**	-4.55	< 0.001	**38.46**	-2.94	0.003
	A	8	62.75	68.75	0.84	0.398	67.47	0.67	0.503	**36.05**	-3.60	< 0.001			
	B		68.60	60.92	-1.06	0.291	**20.93**	-5.99	< 0.001	**14.47**	-6.39	< 0.001	**14.10**	-6.47	< 0.001
	A	9	53.92	58.75	0.65	0.515	53.01	-0.12	0.902	62.79	1.23	0.220			
	B		50.00	58.62	1.14	0.256	**69.77**	2.62	0.009	64.47	1.85	0.065	**65.38**	1.98	0.048
	A	10	69.61	67.50	-0.30	0.761	**50.60**	-2.62	0.009	**16.28**	-6.80	< 0.001			
	B		79.07	71.26	-1.18	0.236	74.42	-0.72	0.471	**26.32**	-6.35	< 0.001	**44.87**	-4.39	< 0.001
	A	11	54.90	61.25	0.86	0.390	55.42	0.07	0.943	**38.37**	-2.25	0.024			
	B		44.19	49.43	0.69	0.490	51.16	0.92	0.360	**14.47**	-3.94	< 0.001	39.74	-0.58	0.565
	A	12	70.59	68.75	-0.27	0.789	73.49	0.44	0.662	66.28	-0.63	0.526			
	B		56.98	44.83	-1.59	0.111	**37.21**	-2.58	0.010	**31.58**	-3.20	0.001	**24.36**	-4.13	< 0.001
Comparison Tasks	*A*	14	58.82	52.50	-0.85	0.394	45.78	-1.76	0.078	**33.72**	-3.39	< 0.001			
	B		63.95	**48.28**	-2.07	0.039	**19.76**	-5.61	< 0.001	**15.79**	-5.81	< 0.001	**42.31**	-2.75	0.006
	A	15	64.71	**47.50**	-2.32	0.021	53.01	-1.61	0.108	**20.93**	-5.75	< 0.001			
	B		73.26	63.22	-1.41	0.158	80.23	1.08	0.280	**52.63**	-2.69	0.007	**53.85**	-2.56	0.010
	A	16	59.80	61.25	0.20	0.843	**34.94**	-3.33	< 0.001	**19.77**	-5.32	< 0.001			
	B		56.98	**74.71**	2.43	0.015	**77.91**	2.89	0.004	46.05	-1.39	0.166	**24.36**	-4.13	< 0.001

Table 4.7: Percentage of participants who answered each question correctly, with statistical information. For this analysis two separate logistic regressions were performed, a 1x4 for Group A, and 1x5 for Group B. Group A represents participants who saw Policies 1 and 2, Group B, participants who saw Policies 3 and 4. Logistic regressions were performed against the standardized table, with z and p values reported above. Percentages in bold indicate statistical differences ($p < 0.05$).

PART II

SMARTPHONE APPLICATION PRIVACY

APPLICATIONS AND THE ANDROID MARKET

RELATED WORK

5

Smartphones allow users to install a variety of applications from different sources, giving the developers of these applications access to a range of personal information. In this chapter we provide background on the Android Market, now Google Play, and the global rise of smartphones. We also outline research on the Android operating system and Android security, as well as on users' expectations with smartphones.

In the past six years Android and iOS, the two now-largest smartphone operating systems, have changed phones from devices with which to call others into true pocket computers. To make this change possible Google, Apple, and others have created operating systems designed around applications. Borrowed, but modified from the desktop model, smartphone applications, or commonly apps, are often small, task-focused executables which users select and install from OS-run application markets.

With access to hundreds of thousands of applications from a diverse and global set of developers come privacy concerns with regards to data permissions. This is only heightened by the amount and variety of personal and sensitive data stored on these modern smartphones. Smartphones by nature store unique cell and phone identifiers, phone contact lists, location data, usage data, and communications information. With use and applications installed, this data can

further include email communications, photographs, social network access, even financial and medical information. And with each new application a user considers, she must determine to grant that developer with access to her data.

Below we will explore the research that has been done to understand the security model of the Android operating system, the current permissions model, and users' expectations regarding their phones. While Apple's iOS, Microsoft's Windows Phone, and BlackBerry's BlackBerry OS each have their own set of interesting properties and levels of user control, we will focus on Android due to its historically more detailed permissions system and its large user base.

5.1 ANDROID AS A MAJOR SMARTPHONE PROVIDER

Since the launch of the first Android phone in fall 2008 the rise of the platform has been spectacular. Android phones accounted for over half of all smartphone sales as of Q3 2011 [45]. With each smartphone sold, more users are downloading applications from the Android Market. As of May 2011, Google reported that over 200,000 applications were available in the Android Market and that those applications had been installed 4.5 billion times in total [8]. As of May 2012, Google has now reported over 15 billion downloads, and over 500,000 applications, with both of these numbers continuing to increase [74].

Applications are not pre-screened, instead users are given the opportunity to decide from all submitted applications which software to install on their phone. Android app rating and recommendation site AppBrain reports that 33 percent of the Applications in the Android Market are rated as "low quality."[1] Additionally, a 2011 Juniper Networks report found "a 472% increase in Android malware samples since July 2011 [to November 2011]" [81]. Similar studies from McAfee [67], Kaspersky Lab [80], and Symantec are all reporting continued exploits. F-Secure's recent Mobile Threat Report (Q4 2012) attributed 79% of all mobile threats in 2012 to Android (up from just 11% in 2010) [32].

Juniper attributes this rise in Android malware to the ease of posting Android applications to the market, as they state: "all you need is a developer account, that is relatively easy to anonymize, $25 and you can post your applications. With no upfront review process, no one checking to see that your application does what it says..." [81].

[1] http://www.appbrain.com/stats/number-of-android-apps

While some believe this openness is harmful to users, Google has promoted it. In one of Google's many tributes to openness, Senior Vice President of Product Management, Jonathan Rosenberg wrote, "At Google we believe that open systems win. They lead to more innovation, value, and freedom of choice for consumers, and a vibrant, profitable, and competitive ecosystem for businesses" [96]. As such, there has been no certification process for Android developers, nor pre-review of applications before they enter the Android Market, though applications reported as malicious have been later removed.

In February 2012, Google's VP of Engineering for Android announced that Google had internally developed a malware blocking tool codenamed Bouncer. He went on to announce that Bouncer had been checking "for malicious apps in Market for a while now," and as a result malware was declining [72]. Our interviews in Chapter 6 were performed before the announcement of Bouncer, while our experiments in Chapter 7 was conducted afterwards. Of course, no malware checker is going to be without error, and there are reports of Bouncer's limitations, such as applications existing in the market for weeks without being noticed [91, 51].

5.2 ANDROID SECURITY RESEARCH

While Android has only existed publicly since 2008, a significant amount of work has been conducted on studying the Android permissions/security model. Much of this research focuses on creating theoretical formalizations of how Android security works or presents improvements to the system security, and is largely out of scope. Enck et al.'s TaintDroid has bridged the gap between system security and user-facing permissions, focusing on analyzing which applications are requesting information through permissions and then sending that data off phone [31].

A follow-up study by Hornyack et al. detailed a method for intercepting these leaked transmissions and replacing them with non-sensitive information [53]. This functionality would allow users post-installation privacy-control. In their investigation they detailed the current permission requests of the top 1100 applications in the Android Market as of November 2010. However, our research, which tests users' understandings of the most common of these permissions, finds users have great difficulty understanding the meaning of these terms. Thus, giving users the ability to limit on a case-by-case basis would likely be ineffective without assistance.

Vidas et al. have also studied how applications request permissions,

finding prevalent "permissions creep," due to "existing developer APIs [which] make it difficult for developers to align their permission requests with application functionality" [111]. Felt et al., in Android Permissions Demystified, attempt to further explain permissions to developers [39]. However, neither of these papers explore end-users understanding of permissions. In our own experiments we find users attempt to rationalize why applications request specific permissions, trying to understand the developers' decisions, even if their understanding of these requests is flawed.

There is also a growing body of research that comes up with novel attack vectors for applications to request more permissions than the users see the application requesting [10]. This work, while interesting, is largely out of scope as we will focus on the permissions the users could expect to have accessed.

Others who have looked at Android permissions have attempted to cluster applications that require similar permissions to simplify the current scheme [9] or have attempted a comparison of modern smartphone permission systems [7]. They find that Android permissions provide the most information to users (compared to other modern smartphone OSs such as Symbian, Windows Phone 7, and iOS), however our interviews show that much of the information provided is not understood.

5.3 ANDROID PERMISSIONS AND PRIVACY RESEARCH

The majority of research done on Android permissions and user expectations has been done by two separate teams at Berkeley. Felt and her colleagues have published a series of papers on the Android permission model, and how users understand it. They show that most users do not pay attention to the permissions screens at install time (17%) and that only three percent of their surveyed users had a strong comprehension of what the permissions were actually asking for access to [38]. They also performed a large risk-assessment survey of users' attitudes towards possible security and privacy risks, and possible consequences of permission abuses [37], a ranking which assisted with our own feature selection in Chapter 7. Finally, they have a paper detailing other possible methods for asking for permission, with a set of guidelines for presenting these privacy and security decisions to users [36]. We will revisit their suggestions in Chapter 8.

Moving away from permissions, King has explored user expectations across the entire use of their smartphones. This broader

investigation, interviewing both iPhone and Android users highlighted difficulties in recognizing the difference between applications and websites, personal risk assessments of possible privacy faults, and how users select applications in the application marketplaces [62].

Neither King nor Felt has proposed and tested alternative permissions displays, or other ways to help users select applications in the Android Market as we will in Chapter 7.

Others have proposed using crowdsourcing to help facilitate users' understanding of applications permissions [116]. While we believe this approach is not without merit, as an understanding of the current types of permissions is beneficial to users, this still involves active awareness of applications requesting permissions in the install process. Lin et al. have more recently explore how some automation and crowdsourcing can be used to map user's expectations of privacy on Android [71].

EXPLORING ANDROID SMARTPHONE USE AND APPLICATION INSTALLATION

EXPERIMENT

6

We performed 20 semi-structured interviews in two cities to determine whether people read and understand Android permissions screens. While doing this we also explored consumer thoughts on the Android ecosystem as a whole, what they hear about Android online and in the news, how they select applications, and what concerns they have about using smartphones.[1]

6.1 RESEARCH QUESTIONS

1. How do Android users use and perceive their smartphones?
2. How well do Android users understand Android permissions?
3. What do Android users consider when downloading new applications?

[1]Portions of this chapter first appeared as "A Conundrum of Permissions: Installing Applications on an Android Smartphone" [61].

6.2 INTRODUCTION

This chapter serves as an initial investigation into how Android smartphone users use and understand their mobile devices. This research was conducted concurrently with the work of King and Felt mentioned in the last chapter.

While security researchers may have a proclivity to focus exclusively on the questions of privacy and security in application choice, this research was conducted with a slightly broader focus towards the use of an Android phone.

With that in mind, we will be framing this interview study around what we believe are the two relevant security and privacy questions that the Android market requires users to make when reviewing potential applications (or apps) for their device.

1. Do I believe this application will compromise the security and function of my phone if I install it?

2. Do I trust this developer and their partners with access to my personal information?

To answer these questions, users leverage word-of-mouth, market reviews and ratings, the Android permissions display, and a host of other considerations to make decisions that protect their mobile privacy and security.

We conducted a series of 20 semi-structured interviews to better understand how users navigate the Android Market, install and use third-party applications, and comprehend the decisions they make at install time.

The remainder of this chapter will detail several variations of the Android permissions displays, our interview methodology, the demographics and expertise of our participants, and finally a collection of participant responses that qualitatively detail their ability to make decisions in the Android ecosystem.

6.3 ANDROID PERMISSIONS AND DISPLAY

Android application permissions are displayed to users at the time they decide to install any app through the Android Market[2] on the web or on the phone. Apps downloaded from third-party app stores (e.g., onlyAndroid, the Amazon Appstore for Android, etc.) do not necessarily show full permissions on their websites, however upon

2. In March of 2012 Google re-branded the Android Market as Google Play. For the remainder of this thesis we will refer to both as simply the Android Market.

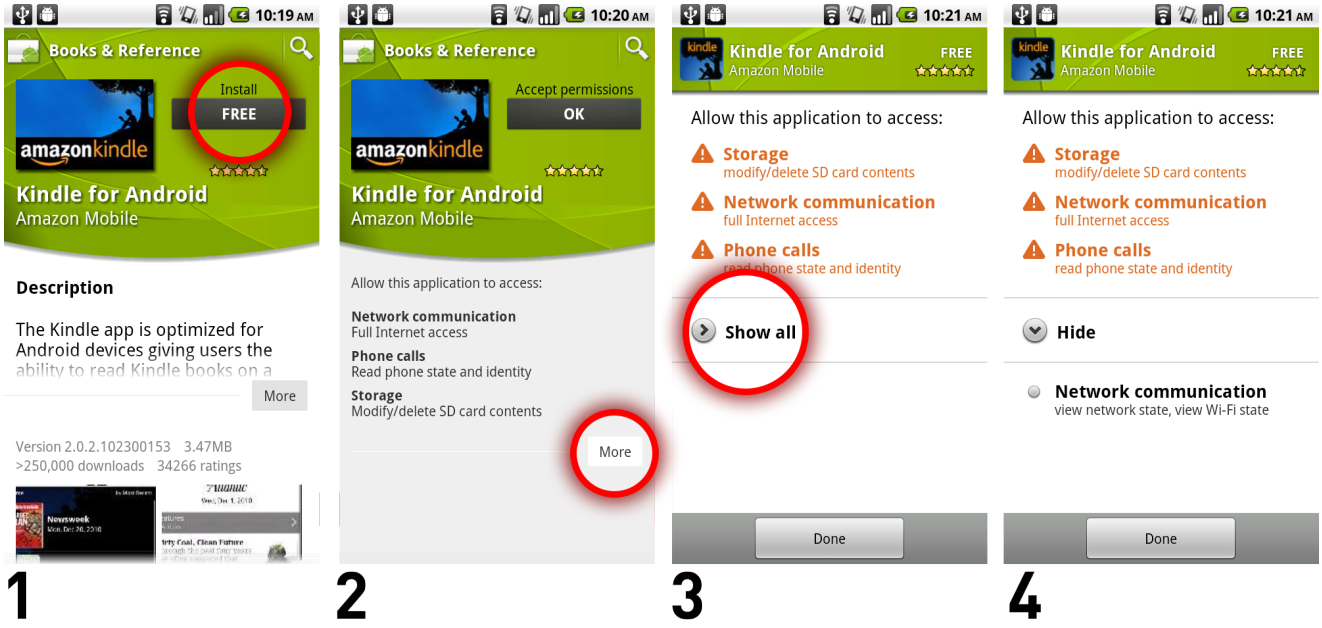


Figure 6.1: The figure above shows the workflow for installing applications and viewing application permissions in earlier versions of the Android Market. Screen 1 shows the Amazon Kindle application as displayed in the Android Market. If a user were to click "FREE," circled in red, they are shown Screen 2, which allows them to Accept permissions and install the application, or to click the "More" button which leads the user to Screens 3 and 4.

installing the application package (APK) the user is presented with a permissions screen variant.

Permissions are shown within the Android Market as detailed in the Figure 6.1. A user browses applications using the view shown in Screen 1. Here there is a truncated description, information about ratings, reviews, screenshots, and much more information about the app, the developer, and related apps. If a user decides to install the app they click the button labeled with the price of the application, here **FREE**. This brings them to Screen 2, where they are given a short list of permissions. If users double tap the **FREE** button on Screen 1, they skip Screen 2 and essentially approve the permissions without viewing the permissions display. Though Screen 2 serves the sole purpose of an interstitial permissions display between the market and a purchase decision, the complete list of permissions is not displayed.

To explore the full application permissions request they would click the **More** expander, bringing them to Screen 3. Here they would see a more complete list of permissions with some permissions shown in red and a **Show all** button, which displays the entire list if toggled. This is a needless multi-step process where a single list would be simpler for users.

At no point in this process is there an explicit way for users to cancel. The only way for users to not install the application after viewing the

permissions is to use the physical back or home buttons on their phone.

The default permissions and groups in the Android SDK are detailed at Android's developer site.[3] The human readable terms displayed in the Android Market are not included in the Android documentation, and have changed with market releases. The definitions shown in the current market screenshots are also not included in the documentation.

6.4 METHODOLOGY

We sought to reach a deeper and more nuanced understanding of how people use Android and how they deal with the issues found in our evaluation of the permissions display above. We conducted semi-structured interviews in Summer 2011 with 20 participants from Pittsburgh and Seattle. The interviews were exploratory in nature, seeking broad understanding of participants' interactions with their smartphones as well as diving deeply into issues surrounding the display of permissions, the safety of the Android Market, and possible harms of information sharing.

We recruited participants through flyers around each city and local Craigslist postings. Each candidate filled out a short pre-survey online before the interview, which allowed us to confirm they did use an Android smartphone. Those participants who opted into the subsequent interview arrived at our labs and completed our consent form allowing us to make an audio recording of their interview. Following the interview participants were given the opportunity to opt-in to share their application information with us, collected through a script running on a local laptop, which we connected their phone to via USB while they watched.

Participants' quotes throughout the remainder of the paper are taken from transcriptions made from the audio recordings of the interviews. Participants were paid $20 for successful completion of the interview, in the form of their choices of Target, Starbucks, or Barnes & Noble gift cards.

[3] http://developer.android.com/reference/android/Manifest.permission.html and http://developer.android.com/reference/android/Manifest.permission_group.html

Participant overview

#	Gender	Age	Occupation	Phone provider	Phone model	OS version	Time Using Android	# Apps downloaded	# Apps really used
1	Female	24	Education	Verizon	LG Ally	I am not sure	1-6 months	1-10	A few 1-5
2	Male	48	Other	Verizon	HTC Incredible	Froyo	1-6 months	11-25	A few 1-5
3	Male	44	Agriculture	T-Mobile	Motorola Cliq	Cupcake	1-2 years	101+	A ton 20+
4	Male	19	Food Service	T-Mobile	Galaxy S	Eclair	1-6 months	11-25	A bunch 6-20
5	Female	45	Legal	Sprint	HTC EVO 4G	Honeycomb	1-6 months	1-10	A bunch 6-20
6	Female	26	Retail	Sprint	Samsung Replenish	I am not sure	1-6 months	1-10	A bunch 6-20
7	Female	34	Engineering	T-Mobile	LG Optimus	Eclair	7 months-1 year	11-25	A few 1-5
8	Male	23	Computers	Verizon	Motorola Droid X	Gingerbread	7 months-1 year	26-100	A ton 20+
9	Female	25	Other	Verizon	Motorola Droid X	I am not sure	Less than 1 month	1-10	A few 1-5
10	Male	32	Engineering	T-Mobile	HTC G2	Eclair	7 months-1 year	11-25	A bunch 6-20
11	Female	21	Entertainment	Sprint	Something Samsung	I am not sure	1-6 months	1-10	A few 1-5
12	Female	22	Other	T-Mobile	HTC MyTouch 4G	I am not sure	7 months-1 year	11-25	A few 1-5
13	Female	21	Don't work	Sprint	HTC Evo Shift	Gingerbread	1-2 years	1-10	A few 1-5
14	Male	20	Real Estate	Verizon	Motorola Droid X	Gingerbread	1-2 years	101+	A bunch 6-20
15	Male	36	Media	Verizon	Motorola Droid 2	Froyo	7 months-1 year	1-10	A few 1-5
16	Male	22	Engineering	Sprint	HTC EVO 4G	Gingerbread	1-6 months	26-100	A bunch 6-20
17	Male	22	Don't work	Verizon	Motorola Droid 2	I am not sure	1-2 years	26-100	A bunch 6-20
18	Female	23	Other	T-Mobile	HTC G2	Gingerbread	More than 2 years	26-100	A bunch 6-20
19	Male	46	Engineering	AT&T	Google Nexus One	Gingerbread	1-2 years	26-100	A bunch 6-20
20	Female	21	Engineering	AT&T	Galaxy S II	Gingerbread	Less than 1 month	1-10	A few 1-5

Table 6.1: Overview of our 20 survey participants. Columns 2-4, list their age, gender, and industry. Columns 5-8 list their phone provider, phone model, Android OS version, and the amount of time they have primarily used Android devices. Columns 9 and 10 show the number of apps they have downloaded and the number they report frequently using. All information is self-reported.

6.5 DEMOGRAPHICS AND SURVEY RESPONSES

Our online survey was completed by 77 participants, 20 of whom completed the lab interview. The remainder of this chapter will discuss solely those 20 users, whose demographic information and survey responses are summarized in Table 6.1. Participants P1-P6 are from Seattle, P7-P20 from Pittsburgh. 10 participants are female, and 10 are male. The ages of our participants range from 19 to 48, with an average of 29. Six of our participants were in tech-related fields, the other fourteen were not. Fourteen of our participants had been using Android for less than a year, five participants reported up to two years of use, and only one reported having used Android for more than two years.

6.6 RESULTS AND DISCUSSION

The following sections detail our findings and participants' responses on various parts of the Android ecosystem. This section is divided into three areas to answer the three research questions posed above. We begin with our participants general impressions on

the Android market. The second area explores users' responses when asked to define ten permissions we asked participants to explain. We conclude with how users self-report their application selection strategies.

Overall, these responses highlight the broad range of often inaccurate knowledge around the Android operating system. From the unclear "human-readable" terms Android provides to users at application install to the often irrelevant coping strategies users have developed.

6.6.1 General Thoughts on Android

User reactions towards Android and their own phones varied widely. While some users loved their Android phones and praised them for openness, a vast array of apps to select from, and generally responsive functionality, others were counting the days until they could switch to a different platform.

One participant who had gotten the phone because her brother had taken advantage of a buy one get one free deal, said, "I want to go backwards so bad, there are way too many things that go on with it. If you have too many things on it, it's just like a computer, it's so slow, I just want to dial a phone number. I got this phone, and a couple months later my fiancee got the same phone, — and we both want to throw them out the window." While her experience was the most negative reaction we saw, others had much more positive things to say. Another said "When I am out with this, I feel just as connected as when I am in front of a computer."

6.6.1.1 Android in the News

We asked participants if they had heard anything about Android phones or Android applications in the news, media, or on the internet. Participants told us about Android's increasing market share, but focused largely on comparisons between iOS and Android. Some had heard one or the other was winning, had better features, was more open, or about new phones for one of the platforms. Most participants could not recall seeing applications advertised in the media. The exception to this rule was of popups that would appear on websites they visited on their phone, recommending they use an app instead of the mobile website.

6.6.1.2 Concern over Malicious Applications

When asked a follow up, to inquire on users' awareness of malicious applications in the Android Market, our participants were largely unaware of any such activity. While a few said they had meant to, or were intending to install anti-virus applications on their phones, most were unconcerned about the threat of malware.

We attribute this lack of concern to two strands we picked up throughout the interviews. The first is that many participants admitted to a lack of trust in new technology. For example, participants reported an unwillingness to do banking from their phone. One participant said "I don't do banking online through my phone because that doesn't seem particularly safe to me.... I prefer an actual desktop for that because I am paranoid." This distrust in something unfamiliar is a coping mechanism we expected, and may cause users to be more cautious with trusting personal data or new applications on their phones.

6.6.1.3 Market Protections

The second part of this lack of concern towards malicious apps shows a deeper misunderstanding of the Android ecosystem. All of our participants, without exception, believed (or hoped) that Android, the entity, was pre-screening applications before entrance into the market. Participants elaborately described the reviews that they thought were taking place, screening not just for viruses or malware, but running usability tests (on actual users!), blocking applications that were too repetitive with other market content, or even screening out applications not enough people would want. Some believed Android was checking for copyright or patent violations, and overall expected Android to be protecting their brand.

Additionally, people were unaware of who was actually running Android. They saw it as a vague entity, that they could not attribute to any specific parent company. Only a few knew and a few more guessed it was Google, others realized they had never stopped to think about that before and were simply unable to attribute the OS to any other company.

6.6.2 Permissions Display Understanding

Half of our participants mentioned the existence of the permissions display before being prompted. When a participant did mention the display, we immediately showed a paper example of one (using the Facebook, Pandora, or Amazon Kindle permissions, Screen 3 of

Figure 6.1). Many participants reported reading, or at least skimming, these displays with some regularity.

Participants were able to identify these screens, recognized them immediately, and occasionally felt very strongly about them. When asked if he read these screens frequently, one such participant said, "Yeah, all the time.... It is just so easy for those apps to do whatever they want, it's a way to protect yourself I guess. Call me paranoid."

Some participants stated that they were not sure how trustworthy the permissions display was. One said of it, "Is it a requirement to be on there [the market] that the software tells you what it is accessing ... Are they required to notify me or not, I don't know."

Unfortunately, most participants do not believe they understand the terms used and have not gone out of their way to learn what they mean. We used a list of ten permissions with the permission group label, in the fashion they would be shown in the permissions display. We asked each user to explain to us their understanding of each term (as if they were explaining it to a relative or friend who was less tech-saavy). Participants reacted to this task with consternation.

Here we present a selection of common, surprising, and strained responses that we received on six of the ten terms we tested.

- **Network communication: full internet access**

  Of the 1100 applications reported on in Hornyack's work [53], Full Internet Access is by far the most requested permission, requested by 941 of the 1100 applications, or 85.5% of those surveyed. Our participants were aware of what the internet is and understood why applications needed it. However how applications have access to it, why they would need to specify it, and how applications would function without it were often unclear.

  - "That you [the user, through the app] can have access to all kinds of websites, even the protected ones." –P1
  - "I would say, this just requires a data plan, and you would need to have internet access." –P6
  - "Any app that needs to get information from somewhere other than that is local on the phone." –P7
  - "For this game to be active, it require internet access, I cannot play it offline." –P11
  - "I would guess that this means, no I don't know. I just assume that it is like taking up data plans. Using stuff with your data plan." –P12

- **Phone calls: read phone state and identity**

  Read phone state and identity is a compound Android permission, allowing access to multiple pieces of functionality. This compound nature seems to lead to participants only correctly anticipating part of the functionality this permission grants. While most of our participants correctly identified functionality related to phone state, the idea that the phone has unique IDs that are also being revealed with this permission was lost on most users. P18 does note a phone ID, but adds an incorrect ability, location. While some applications are requesting this permission to actually detect phone state, many current advertising packages require the identity portion of the permission for uID based tracking.

  - "I would assume it would probably be along the lines of, it knows when my phone is sleeping or in use or in a phone call, and the type of phone" –P2
  - "Phone state whether it is on or off, and identity I would assume it is like my telephone number." –P3
  - "So it knows whether or not I am in the middle of a call? I don't really know what that part [identity] means." –P13
  - "Know where you are, and what phone ID you are on, what type of phone it is." –P18
  - "If you are on the phone maybe it shuts itself off. ... Maybe like your carrier? Hopefully not like who you are." –P19

- **Storage: modify/delete SD card contents**

  Modification and deletion rights themselves were reasonably well understood with users adapting the knowledge they already have, such as computer's memory or thumb drives. However what was stored on the phone itself, compared to the external SD card, was often misunderstood or simply not disambiguated.

  - "That I am about to reach my capacity, or I need to get a new one [SD card]." –P1
  - "Basically, just saving on your memory card or harddrive." –P6
  - "That is for games and things to save your play, store information as needed." –P10

- "It can see what is on my SIM card and on the phone itself." –P13

- **Your location: coarse (network-based) location**

  While we showed participants both types of location that can be collected within Android, participants largely understood that "fine (GPS) location" meant their exact position. It was the coarse location that seemed to confuse more participants. They all understood it was location related, but there was large deviation on how exact that location was.

  - "No, I don't. I haven't the foggiest idea of what that means." –P3
  - "Your network based location, I don't know the difference between the GPS, but basically where you are at." –P6
  - "This is essentially just where your network is located, based on maybe I guess cellphone tower triangulation." –P10
  - "I would guess that this is like the source of your data, like a satellite of some sort." –P12
  - "Is coarse location, does that have anything to do with like, when you have phone service and are in range or roaming?" –P13

- **Your personal information: read contact data**

  Nearly all participants understood that this permission was requesting their address book, or full contact list. Some gave examples of purposes for why this access was needed, citing apps that could use this (P7, P18). A few participants were confused due to the permission group label "your personal information." As a result participants like P11 thought it was reading only data about themselves.

  - "I would think that would mean my contacts list." –P2
  - "Like Facebook, and if it was syncing with contacts." –P7
  - "My phone number." –P8
  - "My personal information can reach them, my name, address, phone number, email address." –P11
  - "Your phone number. They go into your phone, your contacts, and then on Skype they get the number, and he is your friend in your phone. I guess that is what this is." –P18

- **Your accounts: act as an account authenticator**

  This permission was rarely correctly identified (P3, while being unsure, has the right idea), and often described as scary. P12 explicitly said it "freaked" her out. The accounts that participants thought could be "authenticated" or, controlled, were frequently not associated with the application itself, with many participants believing applications that asked for this permission would have very wide-ranging abilities.

  - "Controlling the account? I don't know. I have zero idea." –P2
  - "That I don't like, I don't know what it means, ... my impression is that instead of me being able to authorize something, that application is saying it can." –P3
  - "That freaks me out. What does that mean exactly, cause I am not quite sure." –P12
  - "I dunno is that associated with my T-mobile account?" –P13
  - "I don't know, I guess it is in charge of whatever accounts you open up." –P18

As seen above, for each permission we received answers that we would grade as a misunderstanding. For some of the more obscure permissions, participants simply admitted they didn't know, or gave up. None of our participants correctly understood all ten of the permissions we asked about, and most participants simply repeated the words given in the human readable description, a sign they may not have had complete understanding of the concepts.

Participants asked questions throughout about why applications needed the access they requested. Participants frequently asked the interviewer for examples of applications that requested the permissions we listed, as well as why they were needed. The relationship between the applications and the permissions they requested seemed, without assistance, unknowable.

One participant, when asked if she thought others understood these permissions said, "No. I mean for me to have to think as much, and I have been using these things, and have been sort of a tech-geek for years. Yeah, that's concerning." With Vidas and Felt finding that developers are misunderstanding permissions, and often applying them without need, and self-proclaimed "tech-geeks" finding the terms difficult, common users are left near helpless [39, 111]. The system and terms as they currently stand have not been created or explained for the novice, or even advanced user.

6.6.3 *Application Selection*

While permission information is one vector to assist users in selecting which applications to install, many of our participants reported heavy reliance on star ratings, full text reviews, and word of mouth. These other sources of information were better understood and more trusted.

While reading through the reviews was seen as time-consuming, word of mouth was a trusted way to find high quality applications. One participant recounted his frustrations with searching the store and why he trusted others' opinions: "I feel it is very much a trial and error exercise. And that, I don't know whether that app is a piece of crap or whether it works. So when I know somebody that tells me that this app is good, that really means a lot to me."

Participants also reported hearing about apps, largely of services and products they already used, through advertisements. One participant described his experience with seeing Android app ads, "I have seen magazines and billboards. The phones and the applications. For instance Time Magazine, they have written you can also download the application."

While most of our participants said they do not purchase apps at all, others said in certain cases they would. P6 said, "I try to look for the free ones first, and if I can't find any free ones I will go ahead and buy it."

6.7 CONCLUSION

Users do not understand Android permissions.

Specifically, the human-readable terms displayed before installing an application are at best vague, and at worst confusing, misleading, jargon-filled, and poorly grouped. This lack of understanding makes it difficult for people, from developers to nontechnical users, to make informed decisions when installing new software on their phones. Largely, the permissions are ignored, with participants instead trusting word of mouth, ratings, and Android market reviews.

Users also are largely uninformed about the existence of malware or malicious applications that could be in the Android market. They have difficulty describing the possible harm that could be caused by applications collecting and sharing their personal information. While participants stated they try to find good applications in the market, they believe they are protected by oversight processes which do not exist.

Overall, users are not currently well prepared to make informed privacy and security decisions around installing applications from the Android market.

PRIVACY AS PART OF THE APP DECISION-MAKING PROCESS

EXPERIMENT

7

Understanding that users do not leverage the Android permissions to make application selection decisions, we sought to investigate how we could make permissions and privacy play a part in these decisions. Through an iterative design process and MTurk pilots, we created a short "Privacy Facts" label, which we then tested in 20 in-lab interviews and an online test of 366 participants.[1]

7.1 RESEARCH QUESTIONS

1. What information on the application display screen do Android users consider when downloading new applications?

2. Can we affect users' selection decisions by adding permissions/privacy information to this screen?

3. Do users benefit from, enjoy, and notice this additional information?

[1]Portions of this chapter appear as "Privacy as Part of the App Decision-Making Process" [59].

7.2 INTRODUCTION

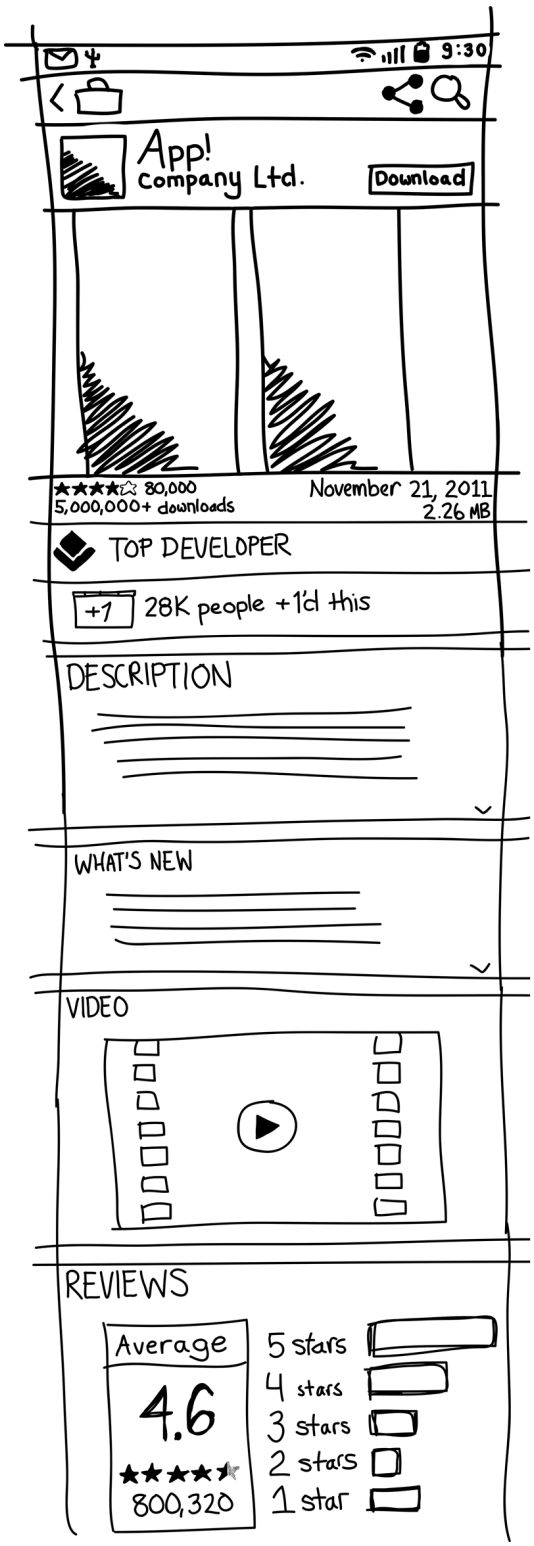
Based on our findings (and the work of others) we have shown the current permissions display does not help users make informed application selection decisions. This final experimental chapter details our design of a short "Privacy Facts" display, which we then tested in a 20-participant lab study and a 366-participant online experiment.

This chapter will first explore the possible design space and describe the concepts we prototyped based on our earlier interviews and evaluation of the current permissions display. The remainder of the chapter will describe three experiments:

1. A series of multiple short tests on MTurk to iterate on our initial design concepts.
2. A 20-participant laboratory interview and application-selection task to receive deep feedback on our proposed design.
3. A 366-participant MTurk study to test if our proposed design can sway users' application selections.

With initial input from MTurkers through a series of micro experiments we refined a single design, which we refer to as the privacy facts checklist. With that design we conducted two studies.

In each study we asked our participants to role-play selecting applications for a friend who has just gotten their first Android phone. Participants were assigned to use either our new privacy checklist or the current permissions display found in the Android market. Our results suggest that our privacy checklist display does affect users' app selection decisions, especially when they are choosing between otherwise similar apps. We also found that both the timing of the privacy information display within the decision-making process and the content of the display may impact the extent to which users pay attention to the information.



Back button, Google Play icon
"Share" icon, magnifying class to search the market

The application icon, name, developer name
Icon for top developer (if applicable)
Download button labeled with cost

Horizontally scrolling view of screenshots

Average rating, number of ratings,
number of downloads, date of last update,
size of application for this phone model

Top developer badge
# of people who have +1'd this application

Text description of the application, with links,
provided by the developer – expands for
descriptions longer than six lines

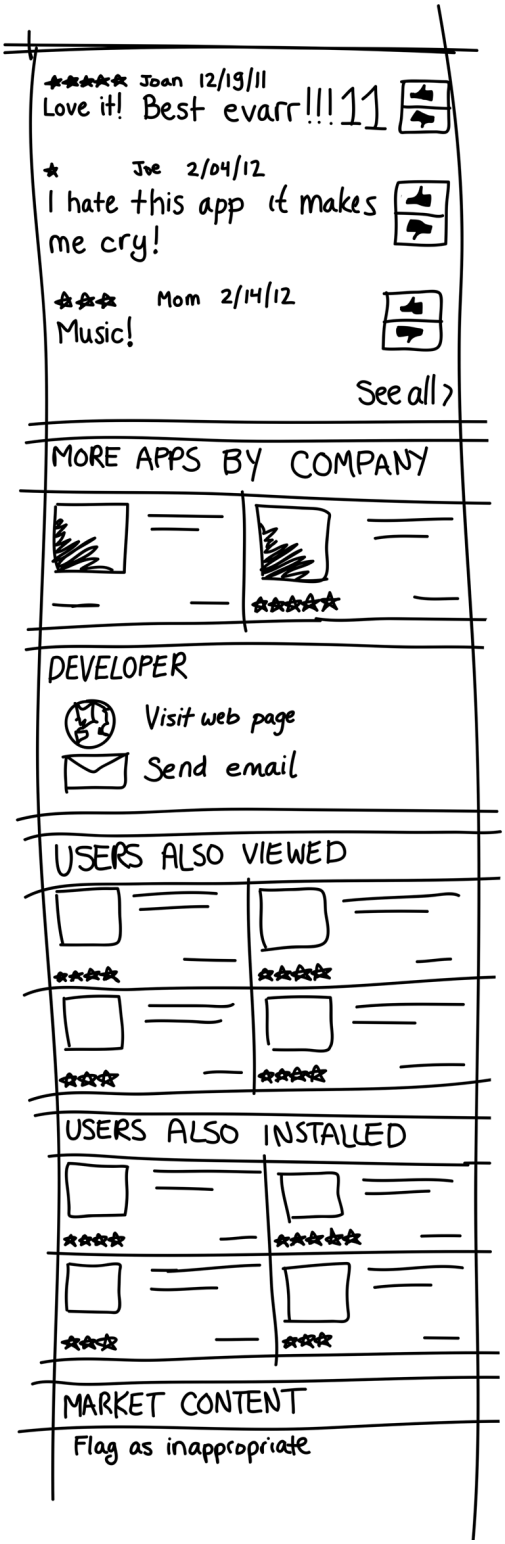
Text description of what is new in the application,
also expandable

Optional video

Reviews section with average review shown by
number and in stars, as well as the number of ratings

Histogram showing the number of 1, 2, 3, 4, and 5
star ratings the application has received

Figure 7.1: "An abstracted view of the top half of the application display screen."



Reviews section continues with a selection of reviews, each shows:

the Google+ name of the reviewer, the date the rating they gave the app, the text of their review, and a way to vote the review up or down

A see all toggle to see more reviews

A selection of other apps by the company, limited to four with a see all toggle

Information about the developer includion optional links to their webpage and a way to send them an email

Other applications that users who looked at this application also viewed, with the ability to see more

Other applicaitons that users who installed this application also installed, with teh ability to see more

A way to flag this app as having inappropriate content for the Google Play store

Figure 7.2: "An abstracted view of the bottom half of the application display screen."

7.3 ANDROID MARKET DESIGN RATIONALE

This section details how the Android Market currently presents privacy information and other information to consumers to help them select new applications to download to their Android smartphone. We detail this space and then discuss a number of possible modifications to make privacy- and security-related information more central to the users' selections.

7.3.1 The Design of the Android Market

Within the Android Market users are presented with a number of ways to search and browse for new applications. Featured applications, top charts, categories, a search tool, and similar application lists each direct users to a common "Application Display Screen" abstracted in Figures 7.1 and 7.2.

This screen provides users with a long list of information about each application. This includes (but is not limited to), a series of navigational items, application information, screenshots, a series of market-assigned labels (top developer, editor's choice), free-test descriptions, a series of reviews, and a series of other types of applications that users may have viewed or chosen.

Given all of the information above, the possibility for expanding several components, and the sheer number of items, the current market application display screen is very long. However, we find users leverage much of this information when making application decisions.

Privacy and security information appears on the above display only when it is mentioned in free-form text by developers or when it appears in text reviews (almost always in a negative context). The entirety of the market-provided, and by extension, system-verified, privacy and security information is on the screen shown after a user has clicked the download button.

This secondary permissions screen (Figure 7.3) again displays the application name, icon, developer, and top developer status icon. This is followed by a very large accept button, which is followed by a list of grouped permissions. Only some permissions are shown initially, followed by a "See all" toggle which expands to display the remainder of the permissions an application requests. Each of these permission groups can be selected to see a pop-up window (Figure 7.5) that contains the definitions for each of the permissions in the selected group. Because there may be several grouped permissions, the pop-ups may have to be scrolled to be completely

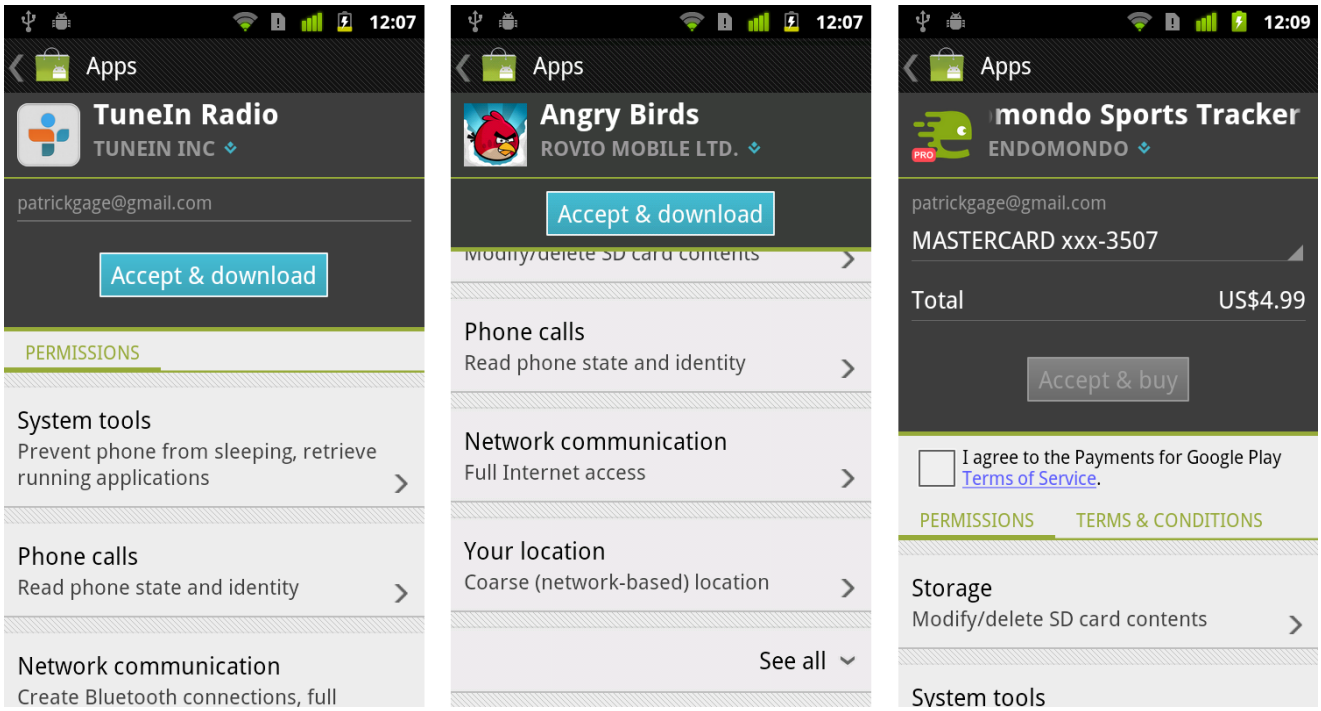


Figure 7.3: The Android permissions display as it exists in September 2012. The left pane shows the page when it first loads to display the permissions of a free application. The middle pane shows the page when scrolling to the bottom where the "See all" toggle can be used to expand the page further. The right pane shows a non-free application. Note in this case only a single permission is visible on first load.

read.

The definitions shown in these pop-up windows, as well as the terms used in the groupings, are not part of the Android developer documents. Only the terms themselves, without definitions, were part of the interviews in the last chapter, due to the more recent introduction of definitions to the market.

7.3.2 *Reasons for modifying the Android Application Display*

We posit that by the time a user selects to move forward by tapping the Download button they have already made their purchase decision. We will see that this is true within our interview study below. For privacy information to be a salient part of the decision process, it must be presented to the user earlier in the process. Privacy information could be included in the long list of other application aspects on the standard application screen. We understand that privacy and permissions information is but one vector of consideration across the entire application-selection process. That said, we propose that it should be placed side by side with the rest of the information on the main application screen to

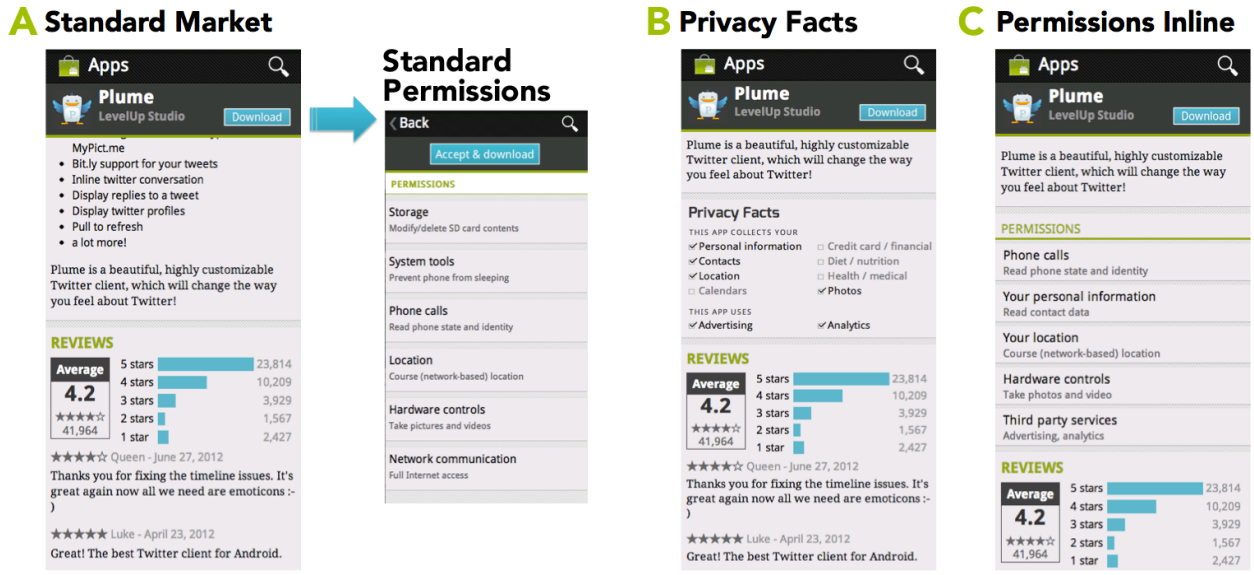


Figure 7.4: The three privacy/permissions display conditions we tested in our experiments.

provide fair consideration and assist users' who choose to use permissions requests to inform their decision.

In an ideation phase, we created a series of possible locations and styles of display for this information. We decided on six realistic and distinct locations for inclusion of privacy information and five different display styles for this information (Figure 7.6).

7.3.3 *Prototype Privacy Facts Design*

The custom privacy display that we tested in our interview and online testing rounds is the Privacy Facts Checklist Display shown in situ (Figure 7.4B. Privacy Facts). The display has several features:

Information— The display has two areas of information. The first with the header "THIS APP COLLECTS YOUR," describes eight types of information the app may collect: Personal information, contacts, location, calendars, credit card/financial, diet/nutrition, health/medical, and photos. The second header specifies "THIS APP USES" and lists advertising and analytics. Each of these ten items has a checkbox next to it, indicating use.

Display Style— The display is 270 pixels tall and the full width of the device (matching all other standard application display sections). For comparison, the rating histogram is 162 pixels tall and the screenshots are the same as our privacy display at 270 pixels.[2] The display has a bold header "Privacy Facts" in a non-Android-standard type.[3] The remainder of the label is presented in the standard

2. There is variation in screenshot size on different Android phone models. The measurements above, and throughout this chapter, are from a Google Nexus One which has an 480x800 pixel display.

3. By using a bold non-standard type we draw attention to this section, and also indicate a sense of authority. The font used is Exo from the Google Font Library.

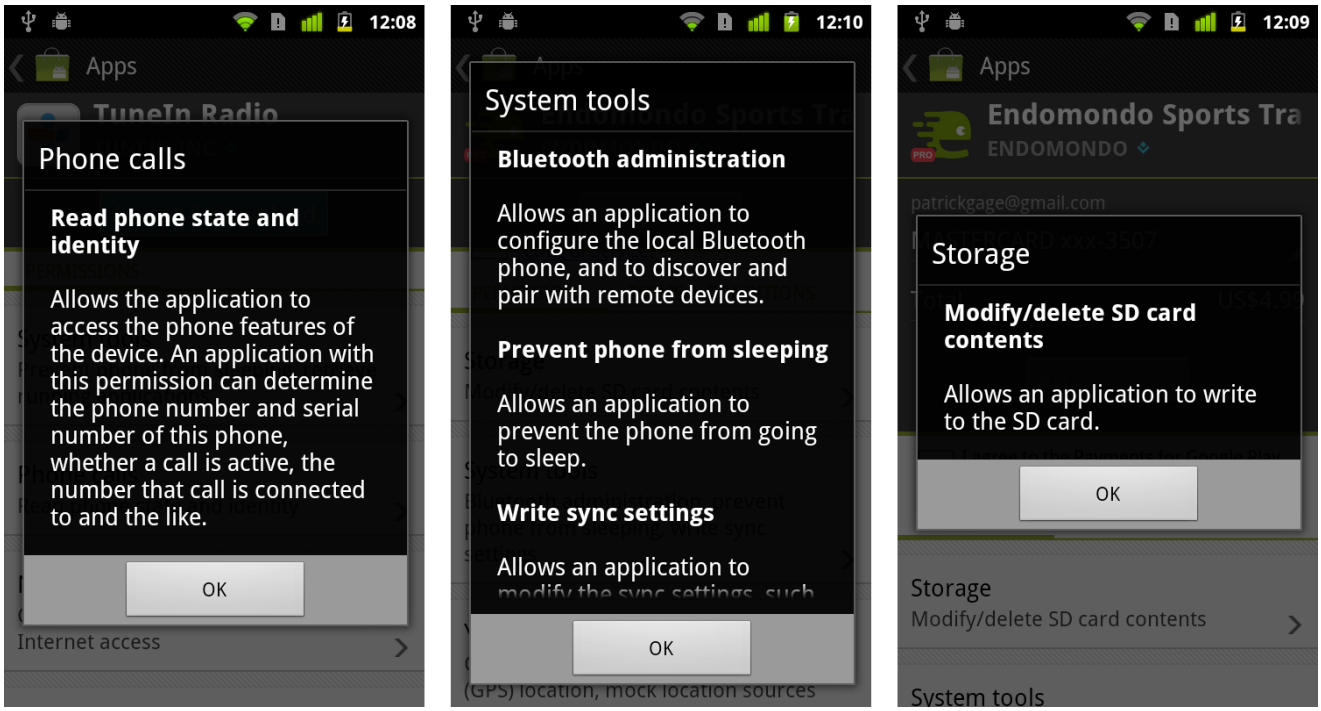


Figure 7.5: These screens show three different pop-ups explaining permissions. The first shows that applications that request a permission called "Read phone state and identity" can in fact see who a user is talking to. The middle pane shows a grouped permission pop-up that requires scrolling to read in full. The third pane shows a short and apparently incomplete definition.

Android Market typeface. The items are each displayed at the standard size, with the headers in capital text in a lighter font color.

Location— The display is shown immediately after the Description section (and Video and What's New sections, if present, which they were not in our studies) and always immediately before the Reviews section. This means when participants first see each app screen there is no visual difference from the market as it is currently displayed. The Privacy Facts section appears below the fold (as it would on most phone models).

Permission mapping— For this display we strayed from the current Android permissions by:

- Including types of information being collected that fall outside of the scope of the current permission model (health information, other financial information),
- Including the use of third-party modules, specifically advertising and analytics
- Removing permissions that are nearly always used (Internet) and those that are irrelevant to most users such as networking protocols and rarely used permissions.

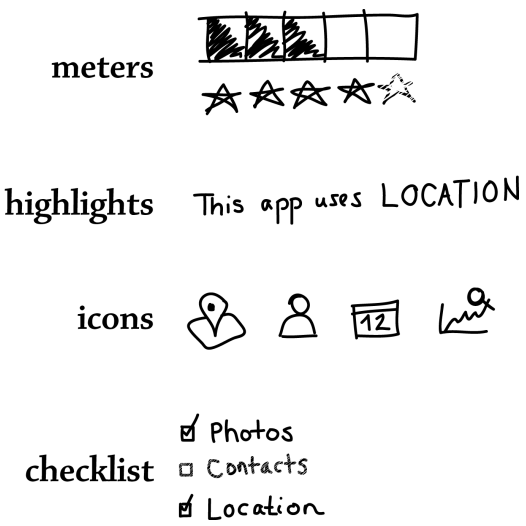


Figure 7.6: We looked at four possible types of information displays to replace or complement the current Android permissions display; those concepts are shown here. We tested meters and checklists in our prototype testing.

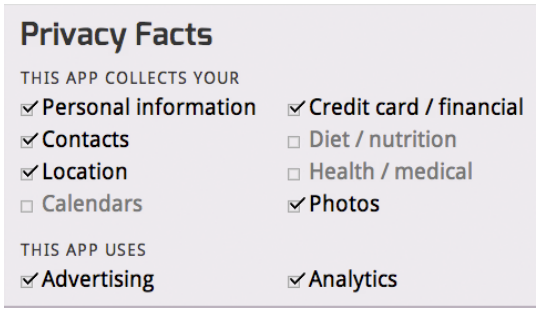


Figure 7.7: The tested Privacy Facts checklist display.

- Including photographs, which are currently accessible to applications.

The final selection of the checklist items we used was strongly influenced by the work of Felt et al. [37] as well as our work in the last chapter. The checklist includes both Android permissions as well as user-provided information. We wanted this display to include both for a more holistic privacy summary. Also, by including an item like photographs, we create a display that is more in line with users' expectations (which universal accessibility of photos is not). A more complex form of this display could include information that explains how these permissions are used, what they are used for, or how frequently they are used.

7.3.4 *Prototyping Round*

This section details a round of iterative MTurk studies. We conducted these short microtests to explore multiple design possibilities before moving to our interview study. We tested two different placements of privacy information: on the main application display screen and on the permissions display screen, the current location. We also tested two of our own design styles: the checklist described above, a privacy meter, the combination of both designs, and the current permissions display.

These microtests were performed for three reasons:

1. To see if the display of privacy information would have any measurable affect.
2. To compare "similar" apps and check for brand/app awareness.
3. To test our custom HTML Android Market.

While we will not detail all of the microtests here, we verified that the applications we were testing were relatively similar in nature. We tested with a set of 18 test applications. We also used these tests to limit the number of possibilities we would need to test in the interview study, and fixed small bugs with our simulated Android Market. We found the Privacy Facts checklist and the Privacy Facts meter (see Figure 7.8) performed similarly, with both having a similar rate of decision change (in early testing approximately a 5% decision shift for every checklist item or 10% bar increase).

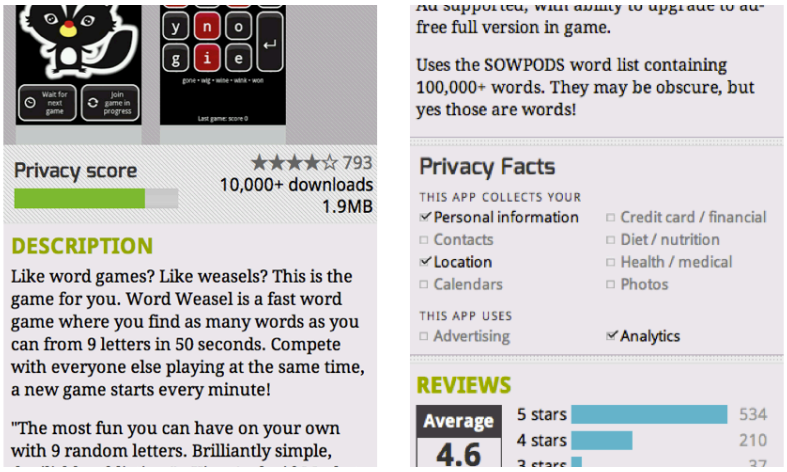


Figure 7.8: The tested Privacy Facts checklist display (on the right) compared to one of the meters we tested in this prototyping round (on the left).

All of the applications we used for testing here and for the remainder of the chapter are real Android applications that could be found and

downloaded in the market. Their names, screenshots, descriptions, features, ratings, and reviews are all authentic. However, we largely picked applications in the 1,000 to 10,000 download range, meaning the applications would not have been seen or used by most participants. All of the applications we tested had 4-star ratings with one exception (Flight Tracker, 3-star) to specifically test rating effects. For all competing application pairs we displayed three text reviews per application, one 2 or 3-star, one 4-star, and one 5-star review.

We also found that not all applications, even after we had adjusted ratings and reviews to be of similar quality, score evenly. While all of our initial word game and nutrition applications selections were relatively comparable, we witnessed unexpected differences in flight tracking and document scanning. In the case of Twitter, one of our earlier choices, twicca, performed so poorly that we removed it and replaced it with Twidroyd. This round allowed us to make fine adjustments to the applications, adjusting ratings, application names, developer names, and other features to try to create balanced app choices.

To test brand, we picked a popular type of application, streaming music. This round showed us that Spotify and Pandora have very high brand recognition, with more than half of our participants recognizing these applications. By continuing to test with Spotify in the next two phases we are able to compare a well recognized application, with a similarly featured but less widely known application, to compare how privacy competes with brand effects.

For the remainder of this chapter we will look at only one custom privacy display, the checklist design, discussed in detail above.

7.4 METHODOLOGY

We ran our experiment in three phases, our MTurk microtests, our 20-participant laboratory exercise and interview study, and an online 366-participant MTurk app comparison survey.

In our studies we asked participants to actively consider how and why they download applications in the market, complete our application-selection task, and then discuss that experience. In both studies, the core of the experiment was an application-selection task using different market designs that vary in how privacy information is presented.

Our study design was based on a similar study run by a team of researchers at Berkeley. The researchers had participants decide

whether to install applications on a computer to see whether people read license agreements at install time. Their users evaluated the software tools as complete packages, based on brand, design, functionality, and also End User License Agreements [48]. Similarly, we seek to understand whether people read the permissions display or our updated privacy facts display when installing software on an Android smartphone, and whether we can manipulate their decisions through improved design and information.

7.4.0.1 Application-selection task

The main task (used in both studies) asked participants to select one application from each of six pairs of applications we presented in our "custom Android market." We presented two applications for each of the six categories (below). All of the applications we used were real applications that could be found and downloaded in the market. Their names, screenshots, descriptions, features, ratings, and reviews were all authentic. However, we picked most applications in the 1,000 to 10,000 download range, such that the applications would not have been seen or used by most participants. We displayed three text reviews per application, one 2- or 3-star, one 4-star, and one 5-star review.

In four of the comparisons we tested applications that were roughly equivalent (Twitter, document scanning, word game, nutrition app). In each of these four cases participants were presented with two applications with different permissions requests, detailed in Table 7.3. In each of these choices one of the applications requested less access to permissions and personal information (low-requesting v. high-requesting).

We also tested two special-case comparisons, to begin to explore the effects of rating and brand. In the flight-tracking comparison, we modified one of the applications (FlightTracker, low-requesting), to have an average rating of 3-stars. All of the other applications in all categories had 4-star average ratings. In the case of streaming music apps, we tested Spotify, a highly-known (shown in pre-tests) application with over 50 million downloads. Nearly all of our participants recognized this application.

7.4.1 Lab Study

To test the privacy facts display, and explore our research question, we conducted a series of semi-structured laboratory exercises in July 2012 with 20 participants in Pittsburgh. This was a between-subjects design. For the main application-selection task ten

participants saw the privacy facts checklist, and the other ten saw the current Android permissions display. We performed exploratory follow-up interviews seeking broad understanding of participants' interactions with their smartphones as well as diving deeply into issues surrounding the display of permissions, understanding of the terms in the checklist/permissions display, the safety of Google Play, and possible harms of information sharing.

We recruited participants through flyers and local Craigslist postings. Each candidate filled out a short pre-survey online before the exercise, which allowed us to confirm they used an Android-enabled smartphone. Those participants who opted into the experiment arrived at our lab and completed our consent form allowing us to make an audio recording of their interview. We performed the study in an on-campus lab and audio recorded the interviews. Participants' quotes throughout the remainder of the paper are taken from transcriptions made from the audio recordings of the interviews.

Participants were assigned randomly to conditions (without any balancing for gender, time using android, technical knowledge, or age). They were paid $20 for successful completion of the interview, in the form of their choice of Target, Starbucks, or Barnes & Noble gift cards.

7.4.2 Exercise and Interview focus

The lab study followed a semi-structured format, outlined here:

- Android introduction: We asked participants basic questions about their Android experience to create a welcoming start and understand their familiarity with the system.
- General new smartphone advice: We asked participants what advice they would give to a hypothetical friend, someone less tech-savvy, who has just gotten a brand new smartphone—what pitfalls they should avoid, what every smartphone user must do, what applications every smartphone user should have.
- Specific new smartphone advice: We then continued the scenario, asking our participants to think about the same friend, but that friend is now looking for six specific applications, and we provided the following text handout:
  - **Word games for killing time —** "I really like word games like Scrabble, but it would be great to have a few things on there for when I need to kill time."

- **Nutrition/Health —** "I keep dieting but an app that helped me keep track of calories would be great."
- **Music —** "I like to listen to music but don't have a large music collection myself."
- **Flight tracking —** "I fly a lot, but I still get a bit anxious and I want to be able to track my flights."
- **Scanning receipts —** "I frequently have to travel for work, and am so bad about keeping all my receipts together. Is there an app that helps me scan in my receipts and save them?"
- **Twitter —** "My friends keep telling me I should use Twitter more, and I do like to follow some celebrities with it, but I don't just want to use the main Twitter app."

We then asked them if they had any specific advice or applications they would recommend for each category. If they weren't sure, we asked them what their strategy for finding an application for that category would be.

- Application-selection task: After verbalizing their suggested applications and strategies we provided them with a Google Nexus One smartphone which we said was their friend's new phone. We showed them a "modified Android Market" and told them that this modified market had two applications from each of the six categories described (12 applications total). These were the two applications their friend was deciding between. We asked them to go take as much time as they needed to decide which was the better choice and to download their selection.

- Post task explanation: After they completed the task we had them explain why they chose the applications they did. We asked a series of questions about what they would have done differently in real life, and if they noticed anything different about the market they used on their friend's phone.

- Android in the news and malicious activity: We then switched gears and had them tell us what they had heard about Android and applications in the news or on the internet. We then asked if they had heard anything about malicious applications and if they had concerns with their information.

- Android permissions and privacy displays: We then drilled down to the privacy and permissions issues, asking if they had noticed the new display or used the current permissions display, depending on condition. We showed them companions of the apps printed on paper, had them define a

series of terms, and asked them if actively thinking about privacy information they would now change their mind.

7.4.3 Online Study

We conducted an online survey, a 366-participant MTurk test of the same application-selection task used in the laboratory study. Because this was performed on MTurk the application-selection task had a more structured survey format, as well as some other methodological differences that will be discussed below in the limitations section. We again used a between-subjects design, but with three conditions. Participants saw one of: the privacy facts checklist (Figure 7.4B); the current android permissions display (Figure 7.4A); or the current android permissions display style and terms presented in the application display screen with additional terms to cover categories from the privacy facts display (Figure 7.4C). In each case they were asked to pick six from the same 12 applications that our participants in the lab study were given, and then were asked to write a short sentence explaining their choice. For successful completion of the survey turkers were paid $0.30.

We used MTurk's user filtering system (95% success required) and required English speakers and Android users. The survey was front loaded with questions about the turker's Android device to discourage users who did not use Android phones. We manually inspected free-response questions to check for participants who were answering randomly, but removed no participants in that stage, only filtering (12) users who had not used the Android market.

7.5 LAB STUDY RESULTS

In this section we detail the results from our lab study. We cover the basic demographics of our participants, their experiences with Android, their advice both general and specific to their hypothetical friend, the results of their application selection, and their post-task interview responses.

7.5.1 Demographics

As shown in Table 7.1, 25% of our 20 participants were male and 75% were female. Participants were between 20 and 44 years old, with an average of 28; 30% were undergraduates. All of our

	Gender	Age	Occupation	Education	Phone Carrier	Phone Model
P1	Female	21	Student	Some college	Verizon	Motorola Droid
P2	Male	21	Student	Some college	Sprint	Motorola Photon
P3	Female	29	Other	Undergraduate	Verizon	Motorola Droid
P4	Female	39	Non-Profit	Undergraduate	T-Mobile	T-Mobile MyTouch 4G Slide
P5	Female	44	Marketing / PR	Undergraduate	Verizon	Pantech Breakout Droid
P6	Female	30	Research / Science	Post-graduate	Verizon	Motorola Droid
P7	Male	43	Other	Some college	AT&T	Motorola Droid
P8	Male	20	Student	Some college	T-Mobile	Motorola Defy
P9	Male	31	Healthcare / Medical	Post-graduate	Verizon	Motorola Droid
P10	Female	23	Research / Science	Undergraduate	Verizon	Samsung Galaxy
A1	Female	20	Student	Some college	Page Plus Pro	Motorola Droid
A2	Female	23	Don't work	Some college	T-Mobile	T-Mobile G2/HTC Desire Z
A3	Female	20	Student	Some college	Sprint	LG Ally / Optimus
A4	Female	28	Student	Undergraduate	Verizon	Samsung Galaxy
A5	Female	24	Student	Undergraduate	Verizon	HTC Rezound
A6	Female	24	Research / Science	Undergraduate	Virgin Mobile	LG Ally / Optimus
A7	Male	21	Student	Some college	Verizon	Motorola Droid
A8	Female	23	Research / Science	Undergraduate	T-Mobile	T-Mobile HTC G2
A9	Female	26	Research / Science	Undergraduate	AT&T	HTC Status
A10	Female	44	Healthcare / Medical	Some college	Verizon	Motorola Droid

Table 7.1: Basic demographics of our lab study participants. Participant numbers beginning with P saw the Privacy Facts display, those with A saw the standard Android system. All the information above was self reported.

participants had downloaded Android applications from the market and were neutral or satisfied with the Google Play experience.

For detailed demographics and Android phone and market information by participant, see Tables 7.1 and 7.2.

7.5.2 Application Selection

The Privacy Facts display appears to have influenced participants in two of the four standard comparisons and in both of the special comparisons. Full selection percentages can be found in the first two columns of Table 7.6 (alongside the online study results). For each participant's selections see Table 7.4.

In two of the four standard comparisons (word game, and Twitter) participants who saw the privacy facts display were, on average, more likely to pick the application that requested fewer permissions. In document scanning, only one participant in each condition did not pick DroidScan Lite (the low-requesting app). In the diet application choice, no participants in the Android condition picked Doc's Diet Diary (the high-requesting app), while three with the

	OS Version	Time Using Android	Satisfaction with the Android Market	Number of Apps Downloaded	Number of Apps Frequently Used
P1	2.3 – Gingerbread	1 – 2 years	4	11–25	1–5
P2	2.3 – Gingerbread	7 months – 1 year	4	101+	6–20
P3	I am not sure	1 – 2 years	4	11–25	6–20
P4	4.1 – Jelly Bean	More than 2 years	5	11–25	20+
P5	2.3 – Gingerbread	7 months – 1 year	5	11–25	6–20
P6	2.3 – Gingerbread	1 – 2 years	4	11–25	6–20
P7	I am not sure	1 – 6 months	3	1–10	None
P8	2.2 – Froyo	1 – 2 years	4	1–10	6–20
P9	4.0 – Ice Cream S.	More than 2 years	3	1–10	1–5
P10	2.2 – Froyo	1 – 2 years	4	11–25	1–5
A1	I am not sure	7 months – 1 year	4	11–25	1–5
A2	2.3 – Gingerbread	More than 2 years	5	1–10	1–5
A3	2.3 – Gingerbread	1 – 2 years	4	26–100	6–20
A4	2.3 – Gingerbread	More than 2 years	4	11–25	6–20
A5	I am not sure	More than 2 years	4	11–25	1–5
A6	I am not sure	1 – 2 years	5	11–25	1–5
A7	I am not sure	1 – 2 years	4	26–100	6–20
A8	2.3 – Gingerbread	1 – 2 years	4	11–25	1–5
A9	3.0 – Honeycomb	1 – 2 years	4	26–100	6–20
A10	4.0 – Ice Cream S.	1 – 2 years	4	26–100	6–20

Table 7.2: Android information on our participants. All the information presented above was self-reported

Privacy Facts display did. In both the two special comparisons more of the participants who saw the privacy facts display picked the low-requesting app.[4]

Participants placed substantial weight on the design and perceived simplicity of using the application. Participants continued to surprise us with ever more idiosyncratic reasons for selecting certain applications. One participant preferred applications with simplistic names, saying "I like to download the apps that have a name that I can easily find. So Calorie Counter, I know where that is gonna be on my phone. I don't have to be like, oh, what is this called."

Participants reported wanting to try the apps out, often saying they would download many and see which was the best (which our study prevented them from doing). One said "And I might try things out and see... I just kind of see how well it works, because some things are more glitchy."

Possible hidden costs also impacted application selection. Several

4. Given the small numbers of participants we did not expect differences to be statistically significant and only the Twitter application choice was significant (Fisher's Exact test, $p = 0.023$, the odds ratio is 11.64).

	Personal Info.	Contacts	Location	Calendars	Credit card/financial	Diet/nutrition	Health/medical	Photos	Advertising	Analytics	Total
Wordoid!	–	–	–	–	–	–	–	–	–	–	0
Word Weasel	X	–	X	–	–	–	–	–	–	X	3
Twidroyd	X	–	–	–	–	–	–	X	–	–	2
Plume	X	X	X	–	–	–	–	X	X	X	6
DroidScan Lite	–	–	–	–	–	–	–	X	–	–	1
Mobile Doc Scan Lite	X	X	–	–	–	–	–	X	–	X	4
Calorie Counter	X	–	–	–	–	–	–	–	–	X	2
Doc's Diet Diary	X	X	X	–	–	X	–	X	–	X	6
Rdio	X	X	–	–	–	–	–	–	–	–	2
Spotify	X	X	X	–	X	–	–	X	X	X	7
Flight Tracker	X	–	X	–	–	–	–	–	–	–	2
iFlights	X	–	X	X	X	–	–	–	–	X	5

Table 7.3: The privacy facts checklist for each application. In each application category, one of the two application requested access to fewer permissions (always shown first).

participants noted that while the music streaming applications were free (as were all the applications we tested), they might have to purchase a subscription, or be unable to access certain functionality after a trial period ended. Participants generally wanted to avoid applications where features would expire or that would require later costs, but more importantly they expected the details of these arrangements to be extremely clear in the descriptions.

7.5.2.1 Android in the news and malicious activity

Most participants reported not seeing much about Android in the news, and most of what they did see was comparisons between Apple's iOS and Android. When we asked about reports of malicious apps, or apps doing unintended things, participants said they had not heard about this. Many believed that it could be hypothetically possible. One participant said "Like, I have wondered, oh could an app be a virus," another "I've heard about viruses, that they can actually shut your computer or phone down. Spyware."

7.5.2.2 *Permissions and Privacy terminology*

To test whether the terms we selected for the Privacy Facts display were understandable, we asked participants to explain what each term meant. While most were very clear, Personal Information and Analytics were the two that participants had the most trouble with. Personal Information answers were often too broad, encompassing things we did not intend. For example, one participant defined it as "That would mean like... interactions within the phone, Gmail, Messaging, Calling different people."

Participants generally preferred the checklist and its terminology. One participant said, "[Privacy Facts is] very straightforward to me. And that is something I noticed, I was thinking, Oh this is cool, is this what they are doing now. That is why I didn't say anything about it. I can immediately go: No, Yes, No, Yes."

Only two participants explicitly mentioned privacy information in their application-selection decisions, both in the Privacy Facts display condition. One participant, said, "If this one is offering the same thing and they want less of your information, I would go with the one that wants less of your information." This comment shows her awareness of the privacy information, but also that the functionality must be matched between apps.

7.5.3 Task time and permission views

Overall, the entire laboratory exercise ranged from 29 minutes to 59 minutes (average 39:53). Participants spent between 3 minutes and 47 seconds to 25 minutes and 6 seconds on the application-selection task. There was no statistically significant difference between conditions (two-tailed t-test, $p = 0.726$), although participants who saw the privacy facts checklist took on average 50 seconds more (11:40 v. 10:51) to complete the task. For complete timing information see Table 7.5.

Across all participants in the Android permissions condition, the permissions screen was used by participants for about half the selection decisions. Four participants decided which applications they would select without ever looking at any permissions screens. Another four participants looked at permissions for all the applications they selected. A6 looked at both Twitter applications permissions, but did not look at the permissions for either of the flight applications. A9 looked at only the permissions for the Twitter application she selected and no other applications.

Across all 31 permission screen views, participants spent between 1

	Word game	Nutrition	Music	Flight tracking	Document scanning	Twitter
P1	Word Weasel	Calorie Counter	Spotify	iFlights	DroidScan Lite	Twidroyd
P2	Wordoid!	Calorie Counter	Spotify	iFlights	DroidScan Lite	Plume
P3	Both	Calorie Counter	Rdio	Flight Tracker	Both	Twidroyd
P4	Wordoid!	Doc's Diet Diary	Spotify	Flight Tracker	DroidScan Lite	Neither
P5	Wordoid!	Calorie Counter	Rdio	Neither	DroidScan Lite	Twidroyd
P6	Wordoid!	Calorie Counter	Spotify	iFlights	DroidScan Lite	Twidroyd
P7	Word Weasel	Doc's Diet Diary	Rdio	Flight Tracker	DroidScan Lite	Twidroyd
P9	Wordoid!	Calorie Counter	Rdio	iFlights	DroidScan Lite	Twidroyd
P9	Wordoid!	Doc's Diet Diary	Spotify	iFlights	DroidScan Lite	Plume
P10	Word Weasel	Calorie Counter	Spotify	Flight Tracker	DroidScan Lite	Twidroyd
	60%	70%	**40%**	**40%**	90%	**70%**
A1	Word Weasel	Calorie Counter	Spotify	iFlights	DroidScan Lite	Plume
A2	Wordoid!	Calorie Counter	Spotify	iFlights	DroidScan Lite	Plume
A3	Wordoid!	Calorie Counter	Spotify	iFlights	DroidScan Lite	Twidroyd
A4	Word Weasel	Calorie Counter	Spotify	iFlights	DroidScan Lite	Plume
A5	Word Weasel	Calorie Counter	Rdio	Flight Tracker	DroidScan Lite	Twidroyd
A6	Wordoid!	Calorie Counter	Rdio	iFlights	DroidScan Lite	Plume
A7	Wordoid!	Calorie Counter	Rdio	Flight Tracker	DroidScan Lite	Plume
A8	Wordoid!	Calorie Counter	Spotify	iFlights	DroidScan Lite	Plume
A9	Word Weasel	Calorie Counter	Spotify	iFlights	DroidScan Lite	Plume
A10	Word Weasel	Calorie Counter	Spotify	iFlights	Mobile Doc Scan Lite	Plume
	50%	**100%**	30%	20%	90%	20%

Table 7.4: Application selections of our participants. The percentages of those participants by condition who selected the application that required fewer permissions are shown below each group, with the better performing condition in bold. Only the Twitter choice difference is statistically significant.

and 11 seconds looking at the Android permissions display. On average they viewed the permissions display for 3.19 seconds (median 2 seconds), including page load time, a minuscule amount compared to time spent on the applications display screen.

7.6 ONLINE STUDY RESULTS

In our online study, the application-selection task was conducted on MTurk through a participant's computer, not a smartphone. Participants saw the applications presented at smartphone size, side-by-side in iframes. Participants selected the application they thought was better for their friend, provided a short text reason, and then rated each of the two presented applications on the likelihood that they would personally acquire it.

With this study, we introduce a third condition, called Permissions Inline. This treatment was designed to separate the location of the

	Time to complete task	Number of permissions screens viewed	Number of seconds spent on permissions
P1	22:06		
P2	3:54		
P3	9:48		
P4	14:51		
P5	9:40		
P6	9:23		
P7	24:20		
P8	25:06		
P9	15:12		
P10	8:29		
Average	11:40		
A1	13:49	6	6, 5, 2, 2, 2, 2
A2	9:50	6	2, 2, 2, 3, 2, 3
A3	8:23	0	–
A4	9:32	0	–
A5	7:03	0	–
A6	13:00	6	4, 2, 10, 5, 11, 1
A7	3:47	0	–
A8	9:24	6	4, 3, 1, 2, 2, 1
A9	16:12	1	9
A10	17:32	6	2, 1, 2, 2, 2, 2
Average	10:51		3.19 seconds

Table 7.5: Timing information for the application-selection task across both conditions and the time spent on permissions screens by participants viewing the permissions display. Participants in the Privacy Facts condition could not look at the standard android permissions displays. The Number of seconds spent on permissions column shows how long a participant looked at a permissions screen each time they did so.

	Lab Study		Online Study						
	Privacy Facts *(n=10)*	Android Display *(n=10)*	*Android Display (n=120)*	Privacy Facts *(n=123)*	Diff. from Android	p-value	Permissions Inline *(n=123)*	Diff. from android	p-value	Inline v. Facts
Wordoid!	60%	50%	40.8%	61.0%			49.6%			
Word Weasel	30%	50%	59.2%	39.0%	**20%**	**0.002**	50.4%	9%	0.198	0.095
Twidroyd	70%	20%	25.0%	52.9%			35.8%			
Plume	30%	80%	75.0%	47.2%	**28%**	**< 0.001**	64.2%	11%	0.051	**0.014**
DroidScan Lite	90%	90%	73.3%	60.2%			62.6%			
M. Doc Scan Lite	0%	10%	26.7%	39.8%	–13%	0.031	37.4%	–11%	0.076	0.784
Calorie Counter	70%	100%	55.8%	73.2%			73.2%			
Doc's Diet Diary	30%	0%	44.2%	26.8%	**17%**	**0.005**	26.8%	**17%**	**0.005**	1
Rdio	40%	30%	17.5%	28.5%			22.8%			
Spotify *(brand)*	60%	70%	82.5%	71.5%	**11%**	**0.048**	77.2%	5%	0.340	0.381
Flight Tracker	40%	20%	40.8%	35.0%			37.4%			
iFlights *(rating)*	50%	80%	59.2%	65.0%	–6%	0.358	62.6%	–3%	0.601	0.791

Table 7.6: Application selections in the laboratory and online studies. The application that requested access to fewer permissions (the privacy-protective choice) is always displayed on top. Statistics for the online study are comparisons to the base Android display. The right-most column shows the significance between the checklist and the inline permissions. Differences in bold, Fisher's Exact. Comparisons with the Android display were planned contrasts. The final comparison between the permissions inline and privacy facts display is Holm-corrected with an adjusted alpha of 0.01667.

privacy information from its format. It showed the standard Android Permissions Display, but positioned on the app display screen (where Privacy Facts is located) rather than in the standard location after the user tapped "Download." This condition tested whether it was only the existence of any privacy information on the application screen that changed behavior, or the checklist format and position.

We used the graphic design of the permissions display from the current Google Play store; however, we modified the labels to present the same information as our Privacy Facts display (including health, nutrition, advertising, and analytics). An example of this is shown in Figure 7.4C.

7.6.1 Demographics

Of our 366 MTurk participants 59% were male and 41% were female (markedly different from our lab study). Our participants were between 18 and 63 years old, with an average of 28. All of our participants had experience downloading Android applications from the market (the 12 who did not were removed from this analysis).

7.6.2 Application selection

Overall the privacy facts display (changed format and position) had a stronger effect on participants application selections than only moving the permissions inline (changed position).

7.6.2.1 Privacy Facts display

In three of the four standard comparisons, significantly more privacy facts participants than Android participants chose the low-requesting app. Only for the document scanner did more participants in the standard Android condition choose the low-requesting app, and this difference was not significant.

For the Twitter choice, nearly three-quarters of the Android display participants chose Plume (high-requesting). One participant captured many of the common reasons for making this choice, reflecting, "Plume has 35,000 more reviews, which suggests to me that this is the more popular, more frequently used application. The description includes a list of everything you can do with the app and those all seem like useful features." However when presented with the privacy facts checklist, the two applications were selected at almost the same rate, with slightly more selecting Twidroyd. Here participants noted and cited the permissions information. One stated, "I picked the one that respects privacy more. The other gets too much personal info." Another participant wrote, "Plume collects too many personal facts."

Between the word games we saw a similar change. Participants using the Android display cited largely reasons of design, color, file size, and ratings, where as participants in the privacy facts checklist condition cited information concerns. One participant who selected Wordoid! said of the application, "It provides a similar experience with similar reviews and does not collect personal information that seem irrelevant to the app's purpose," highlighting that participants will try to understand why an application needs access to certain types of information, and are unhappy when they cannot determine the reason as with these simple word games. One participant made it clear that privacy is a secondary factor in their choice, but in this case the apps were very similar, "It respects your privacy more. The ratings are about the same so privacy wins."

For the special comparisons, rating and brand recognition outweighed privacy. However, even when one of the choices was a well-known brand privacy facts participants were significantly more likely than Android participants to select the relatively-unknown, low-requesting choice. For the flight tracking choice, more

participants chose iFlights (high-requesting) over Flight Tracker. Although participants thought iFlights "sounds like an iPhone port," many believed it had a cleaner UI, but the top reason given was the rating difference. Flight Tracker's 3-stars seems to have outweighed all other factors. For the streaming music choice, Spotify (high-requesting) had much higher brand recognition (although again, both are real services). In the Android permissions display condition over half of the people (66/104) who selected Spotify explicitly stated that they had already heard it was very good or that they or friends use Spotify. One participant said "Spotify is pretty popular and I have never heard of Rdio." Spotify collected much more information than Rdio. but in this case we see that brand information trumps privacy concerns, though there is still a significant shift (11%) in favor of Rdio in the privacy facts condition.

7.6.2.2 Permissions Inline

As shown in Table 7.6, the permissions inline display, while in the same place and often more space-consuming than the checklist, did not have as large an effect on users' decisions. In only one of the four standard comparisons, the nutrition application, was this change significant, and in most cases it underperformed the checklist display (significantly underperforming for twitter apps). This suggests that in addition to moving privacy information to the application display screen, it is important to present that information in a holistic, clear, and simple way if it is to impact users' app selections.

7.6.2.3 Free responses

Across the free-text responses for why applications were selected by participants in the Android display conditions, privacy was only mentioned by one participant, and permissions were mentioned by four others. Across the privacy facts checklist condition privacy was mentioned by 15 participants, and permissions were mentioned by seven more. Information or info were mentioned by 49 people in the privacy facts checklist condition, but by only six participants using the Android display. Based on these responses privacy and personal information seem to have factored more strongly into the decisions of those who saw the privacy facts checklist.

Similar to our lab study, many participants, when directly asked, said they did not notice the privacy facts checklist. Of the 125 participants who were shown the privacy facts checklist, 49 (39.2%) reported in a free-text response having not noticed or paid any

attention to the display. Both those people who did and those who did not notice the display provided reasons for why they ignored it, or believed it was not necessary:

- "I did notice privacy facts, but they weren't the overwhelming reason for any of the decisions."
- "I noticed the Privacy Facts but it really didn't influence me that much. I feel like with social networking it's so much easier to get contacts, photos, or information of someone."
- "It didn't influence my decision even though i noticed it. I tend to pay more attention to ratings and usefulness then anything else."
- "No, not really. It's not the most important factor. I don't keep a bunch of vital personal info on my phone, so no worries. I think people who do are really stupid."

There were also users who found the privacy facts display helpful and made their decisions based on it:

- "I absolutely noticed. I am not a fan of the idea that other people can access my personal info, contacts, location and other things. That is private, and I don't like to give away my information."
- "Yes. I believe the privacy information is helpful. It would only bother me if I saw something that didn't make sense for the app to use. However, I am not terribly concerned about privacy."
- "Yes. It only influenced me if it seemed to be the only thing to distinguish between the two apps."
- "Yeah, I always check that stuff. I want to know exactly what is happening to and with my data from that program when I use it. It was useful though I wish some apps would go into greater detail."
- "Yeah it did. The scanner for instance did not need my personal info, yet the box was checked. This instantly made [me] think the app was a scam. I take everything on the internet with a grain of salt."

Participants who both used and didn't use the display still had misconceptions about companies, sharing information, and the market. Many assumed that all applications collect the same

information. One participant who didn't look at the display said she did not because, "I assume they always say the same thing…."

Participants also continued to believe external forces protect them. One said, "Yes I saw the privacy facts. That didn't really affect my decision as companies are required to protect consumer's information and companies don't really wanna get sued for breach of security so I am not worried about all that." Another stated the continued belief that the market is internally well-regulated, "I think it is trustworthy, I would assume google play [sic] keeps a tight leash on that stuff."

Finally, one participant gave an answer that applies quite broadly, and mirrors work by Staddon et al. [100], "Yes, I noticed the privacy facts but it didn't effect [sic] my decision because I don't really know what the negative impacts of the information they obtain would be." Understanding the potential harm in allowing access to certain types of data remains difficult for consumers both in smartphones and other digital domains.

7.6.3 Self Reported Decision Factors

We also asked our online participants to rank a series of factors in their personal application-selection process from "Not important" to "Very important." The results of this are presented in Figure 7.9. Permissions ranked 8th (of 11), just below two metrics of popularity and just above the size of the application. 80% of participants said ratings and reviews were important or very important, compared with only 43% who said that permissions were important or very important.

This result seems to align with how often participants across our tests tended to ignore permissions.

7.6.4 Limitations

Our short checklist display had items that consumers were able to explain in most cases. Analytics and Personal Information were the most problematic. Participants were generally correct when defining Analytics, but often created more invasive definitions that were not intended. Personal Information was more difficult, as it was too vague and many participants listed other types of data that they then realized were covered by another item on the list. Continuing to refine the terms and types of information that is most important to people will benefit consumers as mobile privacy notices are deployed.

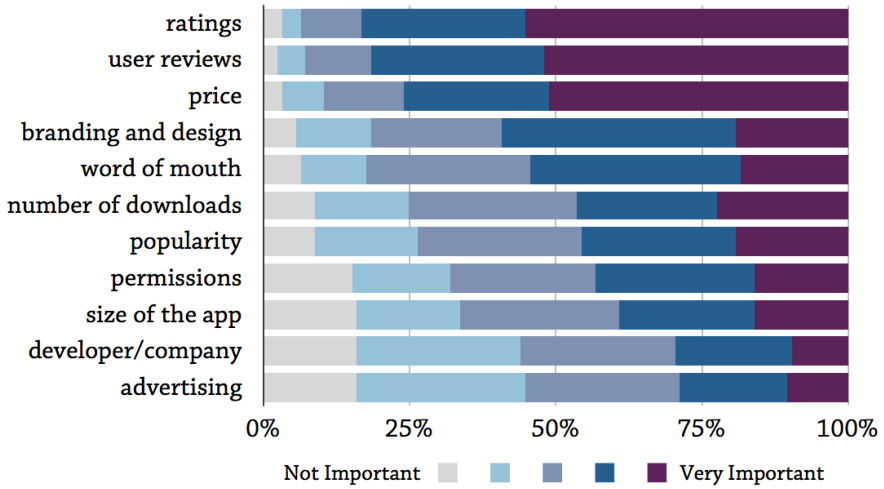


Figure 7.9: A series of factors respondants consider when deciding on applications. Ranked by the number of respondants reporting a 4 or 5, where 5 is "Very important."

One more significant design flaw with the display was that participants do not view permissions displays in the same way as they view privacy policies. They see this information only as items the phone can take, not things that they personally input. While we believe a complete privacy display should cover both user-provided information that is stored (i.e., medical or diet) and automatically collected information like location, this was not explained well by the current design. A possible future design solution is to split the top area of the label into two areas, one that deals with information that can be automatically collected and one that is specifically regarding information collected from the user.

Our lab study has many more female participants than male participants, and due to random condition assignment they were not evenly distributed. We note this as a potential limitation, though our results from the two studies are aligned, and we did not see such a similar gender imbalance in our online study.

Mechanical Turk also has its own set of limitations and biases, which we attempted to counter through a careful survey design. Researchers have shown that MTurk can be used to collect high quality data [25, 54, 63, 106]. While we compared our two survey phases, they did not follow identical methodologies. Our lab study was more realistic, with users using actual cell phones, when on MTurk users saw the applications side by side, and could make direct and visual comparisons. While the reasoning and behavior given seems similar, it is possible that our online survey users had an easier time making decisions, not due to our improved permissions display, but due to the side-by-side display. The only evidence we have to

counter this is the permissions-in-place display did not perform as well as the privacy facts display, implying that the side-by-side display alone is not responsible for all the improvements we saw.

Finally, we tested only 12 applications in the studies described above (and an additional 12 in early pilots). We picked applications that seemed similar, functional, and would be unrecognized, but we would like to expand this work in the future to consider larger application datasets.

7.7 DISCUSSION

We began this chapter with the goal of better understanding how users select Android applications and making privacy and permission information a more salient part of that process.

We found that our participants are using much of the information provided on the application display screen, including some elements we did not include in our own mock market. Yet, we also found users are not using the current permissions display. By moving privacy/permissions information onto the main screen and presenting it clearly we can affect user decisions in cases where applications are similar to begin with. Users mostly appreciated the new privacy facts display, said they would use it to help make their decisions or at least glance at it, and found comparing applications in the market to be a difficult task where better displays would assist them.

7.7.1 What information on the app display screen do Android users consider when downloading new applications?

Most people do not consider permissions when downloading new applications. While this was expected based on other research and our own earlier work, we now have evidence that the permissions are, at least partially, disregarded due to their position in the application-selection process.

Even when instructed to download applications, most users made their decisions without ever pushing the button that would take them to the permissions display. Both our interview participants and our online participants also self-reported that they are aware of the display, but do not look at it. This was confirmed by our interview users who, when they did fully download applications, spent a median time of 2 seconds on the permissions display.

From both the lab and online studies we found that participants continued to report that other characteristics of applications are as important or more important than permissions, including: cost, functionality, design, simplicity, rating, number of ratings, reviews, downloads, size, and others. Continuing to understand how much privacy can compete and offset other aspects is important future work as consumers battle with a crowded and complex market.

The more privacy-hopeful news here is that participants continue to report being familiar with permissions displays and being aware that there are differences between applications. While this may seem unimportant, recognizing that there are privacy differences means creating interfaces that help consumers identify and compare differences could be a path to benefiting users who want to make privacy preserving decisions.

Relatedly, it appears that the current Android application display screen is not well suited to the task of comparing applications at a more general level. Currently, Android consumers browsing for an application search for a term, and then examine, in detail, several of the top, often free, applications. This examination is done sequentially, without an interface for comparison, so all of this information must be retained by the user.

Several participants remarked that the printed, side-by-side comparisons of applications that we showed them at the end of the interview were more useful than the on phone display. While these displays did help them compare the Privacy Facts label, they also found having a side-by-side comparison of reviews, ratings, and features was more useful than browsing on the phone. As with the online privacy policies, we only compared two applications at any time. Developing standardized display formats and interfaces designed for comparison should open up the possibility for multiple application comparisons.

7.7.2 Can we affect users' decisions by adding permissions/privacy information to this screen?

The short answer, is yes—the privacy information on the application display screen affected user behavior. From the microtesting of our prototypes, to the interview responses, to our online test we saw behavioral differences as well as differences in quality and tone of responses relating to private information.

In our online testing we found that having a privacy meter or a checklist would in cases of equivalent applications change user application selection. We did also see that even relatively sizable

privacy differences cannot outweigh some other factors such as very popular applications (some change, but not significant) or differences in average ratings (3-star versus 4-star applications, almost no change at all).

In our interview studies some participants told us they cared more about privacy, but then still selected more invasive applications because of features, design, and brand.

All of our participants had never seen a privacy facts display before, but were immediately able to make comparisons when specifically instructed to do so after the selection task. Most participants, after making these comparisons, did say they would select the alternative, less privacy invasive applications. However, some simply did not believe privacy information was important or relevant to their decision. Some said it would depend on how much their friend (as part of the role-play) cared about his or her own privacy.

These results are similar to those seen in other labeling efforts. Consumers who care more about privacy, whether they have had a credit card stolen or have started receiving spam text-messages, are more likely to take advantage of labeling information. Even if the impact is not drastic, we see the privacy information on the main screen having an effect on selection behavior.

7.7.3 Do users enjoy, notice, and trust permissions information?

Participants in our studies reported being familiar with permissions displays and being aware that there are differences between applications. While this may seem unimportant or obvious, leveraging the awareness of privacy differences means creating interfaces, like checklists, that help consumers identify and compare differences should benefit users who want to make privacy-preserving decisions.

The terms on the current Android Permissions Display remain difficult to understand and participants believed that there was little they could do as most of their information was already exposed. Participants reported that they did not, in most cases, read the information in the displays, and they did not select the permission groupings to see more details or try to better understand the terms. Even when the display was moved to the main screen, it does not have the impact of the privacy facts display.

Participants continued to report not being concerned with data sharing generally, partially due to a belief that companies are following laws and a strong belief that Android/Google is watching

out for their safety as a consumer. While this is accurate in a very general sense, the specifics are quite far off from reality. Correcting the ubiquitous idea of Google Play as a safe, protected marketplace, is important if consumers are to protect themselves through understanding privacy and security in their decision-making process.

Any effort into standardizing a privacy display for smartphones will need to have an educational component associated with it, and while we did not focus on consumer education there is a vast amount of room for improvement. Other studies have highlighted that smartphone users cannot explain the difference between an application and a website [62]. The last chapter shows that most of the permissions terms are not understood by users, and the state of further explanation through definitions is still dismal.

When asked why an application was collecting a type of information, participants most often stated they did not know, but would occasionally venture possibilities. All of our lab study participants wanted to better understand why applications required the permissions they did.

Finally, participants overwhelmingly trusted the application in both the privacy facts display and the permissions display. The question of trusting the information was one most had never considered, and actually gave some participants pause as they realized for the first time that this information might not be accurate. Again, users believe this information is correct, is being verified, and will assume they misunderstand something before they would believe the displays are incorrect. Mistakes in the permissions are not recognized, even when directly discussed. Users will assume they themselves are wrong, not the policy.

PART III

CONCLUSION

DESIGNING PRIVACY NOTICES

8

With our online and Android privacy labels designed and initially verified, we conclude by drawing a series of design trends with the aim of assisting other designers, developers, engineers, and policy makers who are creating privacy information displays.

From our own work—and the designs and interviews of others working in this space—we outline a series of design trends. While these trends are not, at this point, verified as design patterns, we believe they may be of use to those getting started in the space. Additionally, we believe there is benefit in opening them to comment from the community. They cover not just the formatting of privacy notices, but consumer expectations, workflow, and comprehensibility as well as graphic and information design.

8.1 BE AWARE OF EXPECTATIONS

Without fail, consumer expectations and awareness plague efforts to educate and inform consumers about privacy.

Responsible designers, engineers, and policy makers would do well to seek to understand consumer expectations in their specific domain. People's privacy concerns will often be contextual and situational [84, 83].

In the case of online privacy policies we saw that consumers believed that all companies are doing the same things and believed they had no choices with their information, a finding that repeatedly comes up in privacy research [108, 64, 6, 29]. However, in the smartphone arena, consumers did not strongly hold this same belief; they knew that they were granting different applications access to different pieces of information (confirmed by King [62]). This means that we have two different, yet related, problem areas. In online formats we

need to better highlight differences and choice, while in the Android market we would be better served by highlighting more sensitive permissions that consumers can then link to the specific application they are selecting.

We also saw a large difference in the way brand and corporate trust plays into decisions in the two domains we studied. With online privacy policies, consumers are nearly always sharing their personal information with companies they choose based on pre-established relationships and a high level of trust. Even high levels of privacy risk, false advertisement, and security breaches from these companies seem to generally result in little to no active response from consumers, who have invested in these websites. This means that an online privacy notice should not necessarily focus on helping a user make a decision of whether or not to use the website, but could rather highlight the choices a user of that website has regarding optional data collection, third-party sharing, and other services.

On Android the landscape is very different, and application trust is largely created not by brand, but by reviews and ratings. Here, as we showed in the last chapter, privacy can be used to sway application selection. Future work should continue to leverage privacy information in this context at the point of decision, rather than attempting to limit certain types of access to already installed applications.

Without understanding the user expectations in these domains we would not have been able to create privacy notices that leverage these expectations into actionable displays. In other domains, the expectations may again differ from our findings. To design successful notices, some understanding of current consumer knowledge is mandatory.

While consumer action, interpretation, and enjoyability have been a focus of much of our work, we want to clarify that the goal of these (and other) privacy notices, is not to blindly lead everyone to select the most privacy-preserving options possible and cease all information sharing. Substantial work has been done exploring differences between consumer perspectives on privacy, creating clusters based on their attitudes, and understanding how consumers respond to privacy in different situations [112, 66, 108]. Even in areas unrelated to privacy, we see that most labeling standardization efforts are not created to force consumer decisions, but inform.

Privacy notices are best leveraged in domains where consumers do have the freedom to select both more- and less-privacy preserving options. And even if deployed widely we expect many consumers to be unaffected by these displays. Brand, functionality, trust, opinion,

timing, price, and design, will often outweigh privacy. The notices should inform, providing accurate and trusted details in situations where there is elevated concern. Additionally, these notices will be leveraged by consumer watchdogs, journalists, and policymakers, who often have an active desire to support consumer interests.

8.2 PLACEMENT IN THE DECISION PROCESS

If we review the two domains we have covered here, we see that both are, in practice, presenting their privacy notices outside of a user's decision and sharing processes. In the case of the Android market, the privacy information is shown after the user has selected the button displaying the cost of the application. While the user has not fully committed to an application at this point, and can, with some difficulty, abort the operation, in practice the decision has already been made. We tested moving this information and found it did change both how users considered their decision and the actual selections they made, showing that this position is important.

In the case of online privacy policies, we are accustomed to "Privacy" or "Privacy policy" links appear on every page of the website. Sites often have a link and possibly a checkbox confirming agreement when a user signs up with a site. Yet rarely are consumers ever directed to the actual information presented in the privacy policy, especially not at the moment they release personal information to a site. We did not test alternative processes that would directly show this information, as has been done by others looking at EULA and Terms of Service information [48, 56]. In this case, the information must be drastically shortened or summarized to fit into such a process, if there is to be an actual expectation of the user reading the content.

A "teachable moment," can be defined as a point where inserting educational content is more beneficial due to the situation or context the user was just placed in. The concept dates back to at least 1893 [22] but has recently been most leveraged by the health and education communities [76, 101]. It has also been used to educate people about phishing [65]. Other work has looked at the same problem of displaying EULA and privacy information and "the effects of timing" [49, 30].

Again, neither domain we looked at here has attempted to provide the information in such a moment, in a systematic way. Android permissions are listed and agreed to at install time; online privacy policies rarely encourage complete reading. On Apple's iOS granting applications access to some types of information occurs through

pop-up notifications when the application needs the information. While this is temporally closer to when the decision should be made, no explanation for why the permission is requested is given. Users must intuit this from whatever they were just doing, which in many cases is simply launching the application.

With most information available to Android developers as soon as the application is installed on the system, the current Android architecture does seem to call for some privacy education and true consent before install-time. This also makes the common practice of installing an application to test it out, which may seem like a reasonable consumer strategy much more dangerous than users expect. We believe the best solution for the current Android architecture is a blend of permissions explained at install-time, with other more invasive permissions blocked until a later explicit grant.

For a more detailed discussion of how permissions could be better distributed without creating user fatigue see Felt et al. [36].

In the case of online privacy policies, the teachable moment again does not necessarily come at sign-up, where possibly only a small amount of information is being given, but comes throughout a user's natural interactions with the website. When providing credit card, billing, or shipping information, a portion of the screen can remind them what the limits on the use of that information will be, who it will be shared with, how long it will be retained, etc. Then when a user is later posting a review or a photo the details concerning who will be able to view that information, how it will be associated with the user, and more can be explained separately. This removes the burden of having to read a single, long document covering all the possible site interactions, and replaces it with information given as the situation demands.

8.3 UNDERSTANDABILITY

In Chapter 6, we saw that the permissions terminology was largely not understood by users of Android smartphones. Our users thought permissions gave access to more, less, or different information than they did, and inconsistently, and had no idea what many of the permissions meant (verified by Felt et al. [38]).

However, researchers studying privacy notices have known for at least a decade that the terms used in privacy notices are often not well understood by users [55, 78]. This is not because users are uneducated or unteachable, but because the notices tend to be written to protect companies from legal action, at a difficult readability level, with jargon and technical terms. Across these

domains, we see terms created by lawyers and developers simply do not resonate with actual users.

Taking an iterative approach that tests terminology with surveys and focus groups allows for terms to be refined and simplified. This is the approach we have taken in our work, and mirrors standards creation processes [64].

As we continue to create more advanced data sharing and mining techniques, having an awareness of what is and is not understood will allow designers and policy makers to select, reduce, and merge terms. As we have seen in our own efforts to design short notices, merging and hiding terms for the sake of complexity is a necessary part of designing a comprehensible format. While these decisions are difficult to make, they benefit consumers' ability to make decisions. This selection of understandable terms is a process that has not yet happened for online privacy policies and was not performed with enough consumer awareness in the Android market.

8.4 STANDARDIZATION OF TERMS AND FORMAT

Across all the standardized labeling and warning efforts we have seen, term and symbol standardization has been a key part of the design process. For the "Nutrition Facts" panel all of the terms are specified, with regulated short form substitutes, strings of text which must be used when falling under certain requirements (e.g. "Not a significant source of protein"), and ordering of terms is mandated [40]. Additionally, the requirements around how these quantities must be measured—grams, percentage of a daily value are also regulated—to allow consumers a consistent and learnable set of metrics by which to compare their food purchase.

These decisions were not made without debate. Whether or not the daily value percentages are logical, the servings are properly sized, and newer metrics like trans fat should be included continue to be contested [86, 43]. But these metrics allow customers to be presented with information that has been tested and is largely understood. Companies post-NLEA no longer can make diet claims at will nor use standard terminology to make their product sound more beneficial than it is.

While there are obvious benefits of standardized terms which are already widely understood, part of a standardization effort involves defining the terms and educating consumers using these specific definitions. The Nutrition Labeling and Education Act has education

in the title to reinforce that one of their aims was to correct misinformation about health claims and assist users in focusing on different forms of fats, carbohydrates, and vitamins.

We see the same issues in privacy policies and smartphone permissions, but as of yet no body has stepped in to regulate these industries and create a standardized list of terms (though there have been a string of third-party efforts to do so). As a result, companies have created their own, often mediocre, efforts. So while Android has a specific set of application permissions that can be requested, and each of these has a specific display text and pop-up user-definition, these are not included for developers in the Android documentation, as of this writing. Thus, Android has a set of standardized terms, but while they are standard (within the OS version), they are not well defined, and they are also not well understood.

Given this situation, to assist consumer understanding and reap the benefits of standardization, Android should either replace terms with more understandable variants or iterate on their current, limited definitions and begin to educate users on the most relevant of their terms. For true standardization, Android should also reach out to other platform makers: Apple, Microsoft, and Blackberry, to encourage the creation, or use, of shared terms.

We have found standardized formats especially beneficial in comparison tasks. While it is more difficult to tease apart formatting difference—for example, surveys can simply ask users the definitions of words with multiple choice answers, the methodology must change and inevitably become more complex for formatting differences—we believe future work must continue to explore formatting, and how design can be leveraged to make comparison and user choice less burdensome.

One of the failings of the Standard Information Sharing Label discussed earlier was that throughout their entire standard-creation process, thus far, they have spent no time and consideration on the visual design of the format [4]. They acknowledge this on their KickStarter page, “Note: the final design will look MUCH better!” [3]. However, this approach treats the visual design unfairly, as a polish that is added on at the end.

The most successful privacy notices will have formats that are created along with the process of testing terminology, comprehension, awareness and expectations.

8.5 HOLISTIC DESIGN

Finally, we have found value in designing our privacy notices from a holistic standpoint. We define a holistic design as providing:

1. A high-level summary of the most salient points,
2. A clear way to see more details (if they exist) in the context of the whole,
3. All visually present on the screen or paper at once.

Holistic displays have been used to support privacy decision making by Reeder [95] as a primary principle of his Expandable Grid, which was verified in the domains of file permissions and access control. Holistic displays have also been used throughout operations and control design, and advocated for as far back as 1995 to "provide the operator information over a larger span of the process" [102].

The Kleimann Communications Group's financial privacy notice also has this feature, committing all the collection and sharing information to a single page of their label [64].

Showing the entire policy in a single visual space allows users to see:

Each portion of the policy in terms of the whole— This is the same point Strobhar makes, helping his operators understand not just the piece of information they are looking for, but related pieces. This type of view allows someone using an online privacy label to not just end up learning that a company collects phone information for telemarketing, but by scanning down the column, shows them that they may also use web browsing or preference information for telemarketing.

Interactions between elements— Unforeseen interactions between seemingly harmless data elements being collected (or composited later) in tandem continues to cause great consumer fear (see [1, 28]). Our hope is that holistic designs succeed by showing consumers the range of information that will be collected by a single company about them. In future, interactive designs, exploring known "dangerous" combinations could be explored.

What is not used/collected/shared— Finally, nearly all real world privacy notices do not show information that is not used, collected, and shared. The absence of this information is meant to inform consumers that this behavior will not occur. This is not only not in line with consumer expectations, but it makes the act of being more privacy protecting harder to discover. For example, on Android, a customer who wants to know if their location data will not be used

by an application needs to read through all of the permissions listed, and determine if any cover use of location data. If not, they can assume their location data will not be used. However, if they were to do attempt the same thing with the access of photographs on their device, they would find that that permission is never listed, because it is not an Android permission (all applications have access to photographs). They might assume therefore that their photographs will never be collected, when in reality, companies simply do not need to list that data collection.

CONCLUSIONS

9

This thesis shows that it is possible to effectively give consumers informed notice about their privacy decisions. Today, consumers are presented with long, legalistic, jargon-filled, incomplete, and unstandardized notices to make decisions – from which companies they will shop with online to what applications they will download to their smartphone. These notices were not created to serve the purpose of informing the average consumer, but the manner in which they were put in place was due to technical ease or as a legal protection. We explored consumer attitudes towards privacy notices and designed privacy notices that were demonstrably more informative for consumers.

The goal of this work is to explore how improved privacy notices can be created and iteratively improved to help consumers better understand data practices and take more active control of their information.

We tested our own privacy notices in the domains of online privacy policies and smartphone applications in the Android Market. We sought to understand if consumers could better comprehend the information in our formats, if they were able to use our formats quickly, if they enjoyed using our formats, and if they were able to make comparisons and better choices with the information we provided. We found that our standardized, shorter labels were found to be beneficial and ultimately preferred by consumers.

While we have not exhaustively tested our formats—future work should test real-world and long-term use—we believe that they show promise and believe deployment of our labels and variants based on our findings will benefit consumers. We have also quantified how harmful the status quo for privacy notices is today.

Finally, we have provided design suggestions from our domains, synthesized with the work of others, to provide both reflective and actionable material for future designers, developers, and policy

makers who seek to create privacy notices that truly provide users with information on the collection and use of their personal information.

9.0.1 Contributions

This thesis provides several original contributions to benefit those seeking to design and develop privacy notices.

- **A "Nutrition Label" for privacy** — Our proposed standard format for privacy policies is the first tested display of an online privacy policy that benefits consumer understanding, is quick to use, and is liked by users. While other attempts had been previously made, none did significantly better than the status quo online [78].
- **Evaluation of the privacy nutrition label** — Through focus groups, a laboratory study, and a large online test, we showed that other standardized formats, in addition to the nutrition label format, outperformed current text notices. We also found areas where these standardized formats are still weak and could be improved.
- **Application of standardized labels** — We have leveraged a large body of existing work on the design and success of various government standards to build upon those findings towards standardized privacy notices.
- **Continued exploration of consumer attitudes towards privacy** — Attitudes towards privacy inherently change in tandem with the growth and spread of new technologies. Our work continues to explore attitudes towards privacy policies on the web and was some of the earliest work exploring consumer attitudes on smartphone application privacy.
- **Bringing privacy information into the mobile application selection process** — While others have also studied the difficulty consumers have understanding the terms and permissions presented in the Android Market, we proposed a variety of prototypes for displaying privacy and permissions information at a point where it is more relevant to user decisions.
- **Evaluation of a privacy facts checklist for Android applications** — Our work here provides the first alternative privacy and permissions display on Android smartphones, and shows that it leads users to make different, more

privacy-protective decisions than the current permissions display. While further refinement is needed, our interview study and large online test have shown both the power and limitations of privacy information on the Android market.

- **Design suggestions for privacy notices —** Chapter 8 provides five detailed suggestions that we propose should be considered by others creating and deploying privacy notices to consumers. These will serve as a draft of a guide that could later become a series of design patterns to benefit the community, and most of all the consumers, who will benefit from thoughtfully designed, standardized, short, and holistic privacy notices.

BIBLIOGRAPHY

[1] Alessandro Acquisti and Ralph Gross. Predicting social security numbers from public data. Proceedings of the National Academy of Sciences, 2009. http://www.pnas.org/content/early/2009/07/02/0904891106.abstract. Cited on p. 135.

[2] Food Standards Agency. Signpost labeling research, 2005. http://www.food.gov.uk/foodlabelling/signposting/siognpostlabelresearch/. Cited on p. 13.

[3] Joe Andrieu. A standard information sharing label, 2012. http://www.kickstarter.com/projects/joeandrieu/a-standard-information-sharing-label. Cited on pp. 17 and 134.

[4] Joe Andrieu. The standard information sharing label version 0.4, 2012. http://kantarainitiative.org/confluence/download/attachments/58493242/ISWG+Standard+Information+Sharing+Label.Draft+Report.v.0.4.pdf?version=1&modificationDate=1343922915000. Cited on p. 134.

[5] Sheila F. Anthony. Statement of commissioner Sheila F. Anthony concurring in part and dissenting in part, 1999. http://www.ftc.gov/os/1999/07/pt071399anthony.htm. Cited on p. 9.

[6] Annie Anton, Julia B. Earp, and Jessica D. Young. How internet users' privacy concerns have evolved since 2002. In submission, 2009. Cited on p. 129.

[7] K.W.Y. Au, Y.F. Zhou, Z. Huang, P. Gill, and D. Lie. Short paper: a look at smartphone permission models. In Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices (SPSM '11), 2011. Cited on p. 78.

[8] H. Barra. Android: momentum, mobile and more at Google I/O, 2011. http://googleblog.blogspot.com/2011/05/android-momentum-mobile-and-more-at.html. Cited on p. 76.

[9] B. Barrera, H.G. Kayacik, P.C. van Oorschot, and A. Somayaji. A methodology for empirical analysis of permission-based security models and its application to android. In Proceedings of the 17th ACM conference on Computer and communications security (CCS '10), 2010. Cited on p. 78.

[10] David Barrera, Jeremy Clark, Daniel McCarney, and Paul C. van Oorschot. Understanding and improving app installation security mechanisms through empirical analysis of android. In 2nd Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM), 2012. Cited on p. 78.

[11] T.C. Beard, C.A. Nowson, and M.D. Riley. Traffic-light food labels. In Med J Aust., volume 186, page 19, January 2007. Cited on pp. 12 and 13.

[12] Burkey Belser. Designing the food label: Nutrition facts. In greenfield belser, 1994. http://www.greenfieldbelser.com/articles/designing-the-food-label-nutrition-facts. Cited on pp. 11 and 25.

[13] Nick Bilton. Price of facebook privacy? start clicking, May 2010. http://www.nytimes.com/2010/05/13/technology/personaltech/13basics.html. Cited on p. 2.

[14] Ryan Block. Ryan block: Why im quitting instagram, 2012. http://bits.blogs.nytimes.com/2012/12/31/126113/. Cited on p. 2.

[15] E. B. Boyd. Google notifies gmail users of buzz privacy class action settlement, November 2010. http://www.fastcompany.com/1699723/google-settles-buzz-class-action. Cited on p. 2.

[16] P. Buckley and R. Shepherd. Ergonomic factors: The clarity of food labels. In British Food Journal, volume 95, 1993. Cited on p. 25.

[17] Consumer Product Safety Commission. Labeling requirements for toy and game advertisements, 2008. http://cpsc.gov/library/foia/foia08/brief/toygameads.pdf. Cited on p. 13.

[18] European Union Commission. Energy labeling, 1998. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1998:071:0001:0008:EN:PDF. Cited on p. 13.

[19] Josh Constine. How big is facebooks data? 2.5 billion pieces of content and 500+ terabytes ingested every day, 2012.

http://techcrunch.com/2012/08/22/how-big-is-facebooks-data-2-5-billion-pieces-of-content-and-500-terabytes-ingested-every-day/. Cited on p. 1.

[20] Lorrie Faith Cranor. Necessary but not sufficient: Standardized mechanisms for privacy notice and choice. Journal of Telecommunications and High Technology Law, 10(2), 2012. http://ssrn.com/abstract=2184059. Cited on p. 9.

[21] Lorrie Faith Cranor, Serge Egelman, Steve Sheng, Aleecia M. McDonald, and Abdur Chowdhury. P3P Deployment on Websites. Electronic Commerce Research and Applications, 7(3):274–293, Autumn 2008. Cited on pp. 22 and 27.

[22] W. M. Davis. Geography in the schools. The School Review, 1(6):pp. 327–339, 1893. http://www.jstor.org/stable/1074485. Cited on p. 131.

[23] D.M. Dejoy, K.A. Cameron, and L.J. Della. Post-exposure evaluation of warning effectiveness: A review of field studies and population-based research. The Handbook of Warnings, pages 35–48, 2006. Cited on p. 13.

[24] Disconnect. Disconnect db - privacy icons, 2011. https://db.disconnect.me/icons. Cited on p. 17.

[25] J. S. Downs, M. B. Holbrook, Steve Sheng, and Lorrie Faith Cranor. Are your participants gaming the system? screening mechanical turk workers. In ACM CHI, 2010. Cited on p. 121.

[26] Julie S. Downs, George Loewenstein, and Jessica Wisdom. Strategies for Promoting Healthier Food Choices. American Economic Review, 99(2):159–164, 2009. Cited on p. 13.

[27] A.C. Drichoutis, P. Lazaridis, and R.M. Nayga. Consumers' use of nutritional labels. In Academy Marketing Science Review, 2006. Cited on pp. 11, 12, and 58.

[28] Charles Duhigg. How companies learn your secrets, 2012. http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html. Cited on p. 135.

[29] J.B. Earp, A.I. Anton, L. Aiman-Smith, and W.H. Stufflebeam. Examining internet privacy policies within the context of user privacy values. Engineering Management, IEEE Transactions on, 52(2):227–237, May 2005. Cited on p. 129.

[30] Serge Egelman, Janice Tsai, Lorrie Cranor, and Alessandro Acquisti. Timing is Everything? The Effects of Timing and Placement of Online Privacy Indicators. In Proceedings of the

ACM Computer-Human Interaction Conference, New York, NY, USA, 2009. ACM Press. Cited on p. 131.

[31] W. Enck, P. Gilbert, B. Chun, L.P. Cox, J. Jung, P. McDaniel, and A. Sheth. Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. In Proceedings of the 9th USENIX conference on Operating systems design and implementation (OSDI'10), 2010. Cited on p. 77.

[32] F-Secure. Mobile threat report q4 2012, 2013. http://www.f-secure.com/static/doc/labs_global/Research/Mobile%20Threat%20Report%20Q4%202012.pdf. Cited on p. 76.

[33] Federal Trade Commission. Internet site agrees to settle ftc charges of deceptively collecting personal information in agency's first internet privacy case, 1998. http://www.ftc.gov/opa/1998/08/geocitie.shtm. Cited on p. 9.

[34] Federal Trade Commission. Privacy online: Fair information practices in the electronic marketplace, 2000. http://www.ftc.gov/reports/privacy2000/privacy2000.pdf. Cited on pp. 3 and 8.

[35] Federal Trade Commission. Privacy Online: A Report to Congress, June 1998. http://www.ftc.gov/reports/privacy3/toc.htm. Cited on p. 7.

[36] Adrienne Porter Felt, Serge Egelman, Matthew Finifter, Devdatta Akhawe, and David Wagner. How to ask for permission. 2012. http://www.eecs.berkeley.edu/~afelt/howtoaskforpermission.pdf. Cited on pp. 78 and 132.

[37] Adrienne Porter Felt, Serge Egelman, and David Wagner. I've got 99 problems, but vibration ain't one: A survey of smartphone users' concerns. In 2nd Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM), 2012. Cited on pp. 78 and 103.

[38] Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. Android permissions: User attention, comprehension, and behavior. 2012. http://www.cs.berkeley.edu/~afelt/felt-androidpermissions-soups.pdf. Cited on pp. 78 and 132.

[39] A.P. Felt, E. Chin, S. Hanna, D. Song, and D. Wagner. Android permissions demystified. In Proceedings of the 18th ACM

conference on Computer and communications security (CCS '11), 2011. Cited on pp. 78 and 91.

[40] U.S. Food and Drug Administration. Guide to nutrition labeling and education act requirements, 1994. http://www.fda.gov/ICECI/Inspections/InspectionGuides/ucm074948.htm. Cited on pp. 11, 13, 25, and 133.

[41] U.S. Food and Drug Administration. A food labeling guide, 1999. http://vm.cfsan.fda.gov/%7Edms/flg-toc.html. Cited on p. 25.

[42] U.S. Food and Drug Administration. New otc drug facts label, 2002. http://www.fda.gov/FDAC/features/2002/402_otc.html. Cited on p. 13.

[43] U.S. Food and Drug Administration. Trans fat now listed with saturated fat and cholesterol, 2006. http://www.fda.gov/Food/ResourcesForYou/Consumers/NFLPM/ucm274590.htm. Cited on p. 133.

[44] Foursquare. About foursquare, 2012. https://foursquare.com/about/. Cited on p. 1.

[45] Gartner. Gartner says sales of mobile devices grew 5.6 percent in third quarter of 2011; smartphone sales increased 42 percent, 2011. http://www.gartner.com/it/page.jsp?id=1848514. Cited on p. 76.

[46] Guilbert Gates. Facebook privacy: A bewildering tangle of options, May 2010. http://www.nytimes.com/interactive/2010/05/12/business/facebook-privacy.html. Cited on p. 2.

[47] Joshua Gomez, Travis Pinnick, and Ashkan Soltani. Knowprivacy: Privacy coding methodology, 2009. http://knowprivacy.org/policies_methodology.html. Cited on p. 17.

[48] Nathan Good, Rachna Dhamija, Jens Grossklags, David Thaw, Steven Aronowitz, Deirdre Mulligan, and Joseph Konstan. Stopping spyware at the gate: A user study of privacy, notice and spyware. 2005. http://www.cs.umd.edu/~dbthaw/papers/SpywarePaperFinal.pdf. Cited on pp. 106 and 131.

[49] Nathaniel S. Good, Jens Grossklags, Deirdre K. Mulligan, and Joseph A. Konstan. Noticing notice: a large-scale experiment on the timing of software license agreements. In CHI '07:

Proceedings of the SIGCHI conference on Human factors in computing systems, pages 607–616, New York, NY, USA, 2007. ACM. Cited on p. 131.

[50] WELS Regulator (Australian Government). Wels and watermark, 2005. http://www.waterrating.gov.au/compliance.html. Cited on p. 13.

[51] Andy Greenberg. To hide android malware apps from google's 'bouncer', hackers learn its name, friends, and habits, 2012. http://www.forbes.com/sites/andygreenberg/2012/06/04/to-hide-android-malware-from-googles-bouncer-hackers-learn-its-name-friends-and-habits/. Cited on p. 77.

[52] Paulina Haduong, Anthony Tordillos, and Machiste Quintana. Privacy simplified, 2012. http://yale.edu/self/psicons.html. Cited on p. 18.

[53] P. Hornyack, S. Han, J. Jung, S. Schechter, and D. Wetherall. These aren't the droids you're looking for: Retrofitting android to protect data from imperious applications. In Proceedings of the 18th ACM conference on Computer and communications security (CCS '11), 2011. Cited on pp. 77 and 88.

[54] J. J. Horton, D. G. Rand, and R. J. Zeckhauser. The online laboratory: Conducting experiments in a real labor market. In Experimental Economics, 2010. Cited on p. 121.

[55] Carlos Jensen and Colin Potts. Privacy policies as decision-making tools: An evaluation of online privacy notices. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pages 471–478, Vienna, Austria, 2004. Cited on pp. 2, 10, and 132.

[56] Matthew Kay. Techniques and heuristics for improving the visual design of software agreements. Master's thesis, University of Waterloo, 2010. http://hdl.handle.net/10012/5483. Cited on pp. 19 and 131.

[57] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. A "Nutrition Label" for Privacy. In Proceedings of the 2009 Symposium On Usable Privacy and Security (SOUPS), 2009. Cited on p. 21.

[58] 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, CHI '10, pages 1573–1582, New York, NY, USA, 2010. ACM. Cited on p. 49.

[59] Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. Privacy as part of the app decision-making process. In ACM CHI, 2013. Cited on p. 95.

[60] Patrick Gage Kelley, Aleecia McDonald, Robert Reeder, and Lorrie Faith Cranor. P3P expandable grids, 2007. http://cups.cs.cmu.edu/soups/2008/posters/kelley.pdf. Cited on pp. 10 and 24.

[61] PatrickGage Kelley, Sunny Consolvo, LorrieFaith Cranor, Jaeyeon Jung, Norman Sadeh, and David Wetherall. A conundrum of permissions: Installing applications on an android smartphone. In Financial Cryptography and Data Security, volume 7398, pages 68–79. 2012. Cited on p. 81.

[62] Jennifer King. How come I'm allowing strangers to go through my phone?: Smart phones and privacy expectations, 2013. http://jenking.net/mobile/. Cited on pp. 79, 125, and 129.

[63] Aniket Kittur, E. H. Chi, and B. Suh. Crowdsourcing user studies with mechanical turk. In ACM CHI, 2008. Cited on p. 121.

[64] Kleimann Communication Group Inc. Evolution of a prototype financial privacy notice., February 2006. http://www.ftc.gov/privacy/privacyinitiatives/ftcfinalreport060228.pdf. Cited on pp. 9, 14, 15, 26, 34, 129, 133, and 135.

[65] Ponnurangam Kumaraguru. Phishguru: A system for educating users about semantic attacks, 2009. http://reports-archive.adm.cs.cmu.edu/anon/isr2009/CMU-ISR-09-106.pdf. Cited on p. 131.

[66] Ponnurangam Kumaraguru and Lorrie Faith Cranor. Privacy indexes: A survey of westin's studies. Technical Report CMU-ISRI-5-138, Carnegie Mellon University, December 2005. http://reports-archive.adm.cs.cmu.edu/anon/isri2005/CMU-ISRI-05-138.pdf. Cited on p. 130.

[67] McAfee Labs. Mcafee threats report: Third quarter 2011, 2011. http://www.mcafee.com/us/resources/reports/rp-quarterly-threat-q3-2011.pdf. Cited on p. 76.

[68] Pr Lanner. Previewing online terms and conditions: Commonterms alpha proposal, 2012. http://commonterms.net/commonterms_alpha_proposal.pdf. Cited on p. 19.

[69] Pedro Giovanni Leon, Justin Cranshaw, Lorrie Faith Cranor, Jim Graves, Manoj Hastak, Blase Ur, and Guzi Xu. What do online behavioral advertising disclosures communicate to users?, 2012. http://www.cylab.cmu.edu/files/pdfs/tech_reports/CMUCyLab12008.pdf. Cited on p. 18.

[70] Alan Levy and Manoj Hastak. Consumer comprehension of financial privacy notices: A report on the results of the quantitative testing, 2008. http://www.ftc.gov/privacy/privacy\initiatives/Levy-Hastak-Report.pdf. Cited on pp. 15 and 70.

[71] Jialiu Lin, Norman Sadeh, Shahriyar Amini, Janne Lindqvist, Jason I. Hong, 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, UbiComp '12, pages 501–510, New York, NY, USA, 2012. ACM. Cited on p. 79.

[72] Hiroshi Lockheimer. Android and security, 2012. http://googlemobile.blogspot.com/2012/02/android-and-security.html. Cited on p. 77.

[73] Tom Lowenthal and Alex Fowler. Privacy icons v0.2, 2011. https://wiki.mozilla.org/Privacy_Icons_v0.2. Cited on p. 17.

[74] Ingrid Lunden. Google play about to pass 15 billion app downloads? pssht! it did that weeks ago, 2012. http://techcrunch.com/2012/05/07/google-play-about-to-pass-15-billion-downloads-pssht-it-did-that-weeks-ago/. Cited on p. 76.

[75] N. Maubach and J Hoek. The Effect of Alternative Nutrition Information Formats on Consumers' Evaluations of a Children's Breakfast Cereal. Proceedings of the EParternships, Proof and Practice – International Nonprofit and Social Marketing Conference, 2008. Cited on p. 13.

[76] C. M. McBride, K. M. Emmons, and I. M Lipkus. Understanding the potential of teachable moments: the case of smoking cessation. Health Education Research, 18(2):156–170, 2003. Cited on p. 131.

[77] Aleecia McDonald and Lorrie Cranor. The cost of reading privacy policies. In Proceedings of the Technology Policy

Research Conference, September 26–28 2008. Cited on pp. 2 and 10.

[78] Aleecia M. McDonald, Robert W. Reeder, Patrick G. Kelley, and Lorrie F. Cranor. A comparative study of online privacy policies and formats. In Proceedings of 2009 Workshop on Privacy Enhancing Technologies. ACM, 2009. Cited on pp. 2, 10, 40, 132, and 138.

[79] Matthias Mehldau. Iconset for data-privacy declarations v0.1, 2007. http://asset.netzpolitik.org/wp-upload/data-privacy-icons-v01.pdf. Cited on p. 17.

[80] Y. Namestnikov. It threat evolution: Q3 2011, 2011. http://www.securelist.com/en/analysis/204792201/IT_Threat_Evolution_Q3_2011. Cited on p. 76.

[81] Juniper Networks. Mobile malware development continues to rise, android leads the way, 2011. http://globalthreatcenter.com/?p=2492. Cited on p. 76.

[82] Nicholas Carlson. Warning: Google buzz has a huge privacy flaw, Feb. 10, 2010. http://www.businessinsider.com/warning-google-buzz-has-a-huge-privacy-flaw-2010-2. Cited on p. 2.

[83] Christina E. Nippert-Eng. Islands of Privacy. University of Chicago Press, 2010. Cited on p. 129.

[84] Helen F. Nissenbaum. Privacy as contextual integrity. 79, 2004. http://ssrn.com/abstract=534622. Cited on p. 129.

[85] Commonwealth of Australia. The Energy Label, 2007. http://www.energyrating.gov.au. Cited on p. 13.

[86] Peter L. Pellett. Commentary: The r.d.a. controversy revisited. Ecology of Food and Nutrition, 21(4):315–320, 1988. http://www.tandfonline.com/doi/abs/10.1080/03670244.1988.9991045. Cited on p. 133.

[87] Marc Perry. Library of congress, facing privacy concerns, clarifies twitter archive plan, May 7, 2010. http://chronicle.com/blogs/wiredcampus/library-of-congress-facing-privacy-concerns-clarifies-twitter-archive-plan/23818. Cited on p. 2.

[88] Travis Pinnick. Privacy short notice design, 2011. http://www.truste.com/blog/2011/02/17/privacy-short-notice-design/. Cited on p. 16.

[89] Privacy Leadership Initiative. Privacy notices research final results, December 2001; Accessed: December 17, 2007. http://www.ftc.gov/bcp/workshops/glb/supporting/harris%20results.pdf. Cited on p. 9.

[90] PrivacyChoice. Privacychoice - make your policy, 2012. http://privacychoice.org/policymaker. Cited on p. 17.

[91] Fahmida Y. Rashid. Black hat: Researchers find way to "bounce" malware into google app store, 2012. http://www.scmagazine.com/black-hat-researchers-find-way-to-bounce-malware-into-google-app-store/article/252098/. Cited on p. 77.

[92] Aza Raskin. Is a creative commons for privacy possible?, 2010. http://www.azarask.in/blog/post/is-a-creative-commons-for-privacy-possible/. Cited on p. 17.

[93] Aza Raskin. Privacy icons: Alpha release, 2010. http://www.azarask.in/blog/post/privacy-icons/. Cited on p. 17.

[94] Robert Reeder, Lorrie Cranor, Patrick Kelley, and Aleecia McDonald. A user study of the expandable grid applied to P3P privacy policy visualization. In Workshop on Privacy in the Electronic Society, 2008. Cited on pp. 24 and 25.

[95] Robert W. Reeder. Expandable grids: A user interface visualization technique and a policy semantics to support fast, accurate security and privacy policy authoring, 2008. http://www.robreeder.com/pubs/ReederThesis.pdf. Cited on pp. 22 and 135.

[96] J. Rosenberg. The meaning of open, 2011. http://googleblog.blogspot.com/2009/12/meaning-of-open.html. Cited on p. 77.

[97] Hugo Roy, Michiel de Jong, and Jan-Christoph Borchardt. Terms of service; didn't read, 2012. http://tos-dr.info/. Cited on p. 19.

[98] Mary Rundle. International data protection and digital identity management tools, 2006. http://identityproject.lse.ac.uk/mary.pdf. Cited on p. 17.

[99] J.D. Seymore, A. Yaroch Lazarus, M. Serdula, H.M. Blanck, and L.K. Khan. Impact of nutrition environmental interventions on point-of-purchase behavior in adults a review. Preventative Medicine, 29:S108–S136, 2004. Cited on p. 12.

[100] Jessica Staddon, David Huffaker, Larkin Brown, and Aaron Sedley. Are privacy concerns a turn-off? engagement and privacy in social networks. In Symposium on Usable Privacy and Security (SOUPS), 2012. Cited on p. 120.

[101] V J Stevens, H Severson, E Lichtenstein, S J Little, and J Leben. Making the most of a teachable moment: a smokeless-tobacco cessation intervention in the dental office. American Journal of Public Health, 85(2):231–235, 2012/08/23 1995. http://dx.doi.org/10.2105/AJPH.85.2.231. Cited on p. 131.

[102] David A. Strobhar. Evolution of operator decision making. ISA Transactions, 34(4):405 – 409, 1995. http://www.sciencedirect.com/science/article/pii/001905789500033X. Cited on p. 135.

[103] Kara Swisher. Mary meeker explains internet 2012 in 17 minutes: The full d10 interview (video), 2012. http://allthingsd.com/20120612/mary-meeker-explains-internet-2012-in-17-minutes-the-full-d10-interview-video/. Cited on p. 1.

[104] The Center for Information Policy Leadership. Multi-Layered Notices Explained, 2004. http://www.hunton.com/files/tbl_s47Details/-FileUpload265/1303/CIPL-APEC_Notices_White_Paper.pdf. Cited on pp. 16 and 53.

[105] The Center for Information Policy Leadership. Ten steps to develop a multilayered privacy notice, 2005. http://www.hunton.com/files/tbl_s47Details/-FileUpload265/1405/Ten_Steps_whitepaper.pdf. Cited on pp. 16 and 53.

[106] M. Toomim, Travis Kriplean, C. P□rtner, and James Landay. Utility of human-computer interactions: toward a science of preference measurement. In ACM CHI, 2011. Cited on p. 121.

[107] J. Tsai, P. Kelley, L. Cranor, and N. Sadeh. Location-sharing technologies: Privacy risks and controls. In 37th Research Conference on Communication, Information and Internet Policy (TPRC '09), 2009. Cited on p. 2.

[108] Janice Y. Tsai. The impact of salient privacy information on decision-making, 2009. http://repository.cmu.edu/cgi/viewcontent.cgi?article=1000&context=dissertations. Cited on pp. 129 and 130.

[109] J. Turow, L. Feldman, and K. Meltzer. Open to exploitation: American shoppers online and offline. Technical report, Annenberg Public Policy Center of the University of Pennsylvania, 2005. Cited on p. 9.

[110] United States Code. 6803. Disclosure of institution privacy policy, 2008. `http://www.ftc.gov/privacy/glbact/glbsub1.htm#6803`. http://www.law.cornell.edu/uscode/15/usc_sec_15_00006803----000-.html. Cited on p. 14.

[111] Tim Vidas, Nicolas Christin, and Lorrie Faith Cranor. Curbing android permission creep. In W2SP 2011, 2011. Cited on pp. 78 and 91.

[112] Alan Westin. Privacy and Freedom. New York: Atheneum, 1967. Cited on pp. 1, 3, and 130.

[113] Moms with Apps. Privacy icon, 2012. http://momswithapps.com/privacy-icon/. Cited on p. 18.

[114] World Wide Web Consortium. The platform for privacy preferences 1.0 (P3P1.0) specification, 2002. http://www.w3.org/TR/P3P/. Cited on p. 22.

[115] World Wide Web Consortium. The platform for privacy preferences 1.1 (P3P1.1) specification, 2006. http://www.w3.org/TR/P3P11/. Cited on p. 34.

[116] Liu Yang, Nader Boushehrinejadmoradi, Pallab Roy, Vinod Ganapathy, and Liviu Iftode. Enhancing users' comprehension of android permissions. In 2nd Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM), 2012. http://www.cs.rutgers.edu/~vinodg/papers/spsm2012/spsm2012.pdf. Cited on p. 79.

ACKNOWLEDGMENTS

Over the course of the last seven years it has become difficult for me to encourage others to pursue doctoral work. I have seen too many friends and colleagues become embittered against their own research, their advisors, and often all of academia. I can only attribute my own relatively stress-free and truly happy experience to my own quirky mental structure or, more likely, the simply excellent group of people who supported me. You all mean more to me than I can ever properly express.

My family, all the cousins and aunts and uncles that made opening Christmas presents an explosion of wrapping paper, have been supportive, inquiring, and remain confused as to why I wanted to be in school for more than two decades. My grandparents, especially Helen Gage, are irreplaceable: for putting up with my dad, teaching me to make pizza, playing rummy with me, encouraging me to draw, and being there for me as I grew up. Katie, I owe you for mailing cards out to every family member on every special occasion I forgot and still signing my name. Someday I will find all the presents I have accumulated for you from my travels and they still won't be enough to thank you. Dad, I can't say how different I would have turned out growing up in a house that didn't have 15,000 books in it, I am doing my best to emulate that in Albuquerque. Mom, I appreciate you always keeping me focused and grounded, even if I still think that the full time poetry route would have probably been even more exciting.

Carnegie Mellon truly deserves all the praise it receives for producing excellent scholars, researchers, and students. The students and faculty I was able to work with at the CUPS and Mobile Commerce Lab are excellent scientists and became great friends.

Rob Reeder, Serge Egelman, Ponnurangam Kumaraguru, Steve Sheng, Janice Tsai, Kami Vaniea, and Michael Benisch were the best senior graduate students I could have asked for as a clueless first-year. Paul Hankes-Drielsma, Eran Toch, Jonathan Mugan, Robert McGuire, Aleecia McDonald, Cristian Bravo-Lillo, Yang Wang, Rebecca Balebako, Pedro Giovanni Leon, Peter Klemperer, Dave Gordon, Ian Fette, Rich Shay, Saranga Komanduri, Idris Adjerid, Blase Ur, Michelle Mazurek, Justin Cranshaw, and Manya Sleeper all put up with my endless complaints and cynicism at research meetings. I promise to help with graph and table formatting as long as you will have me. I was also able to work with a series of talented undergraduates who did my work better than I could, including: Lucian Cesca, Joanna Bresee, Daniel Rhim, Jerry Feng, Steve Won, Yael Mayer, and Robin Brewer.

The researchers and faculty in my HCI/Privacy/Security community have showed me who I hope to be: Lujo Bauer, Howard Seltman, Jason Hong, Jodi Forlizzi, John Zimmerman, Niki Kittur, Ben Fry, Carlos Guestrin, Osman Khan, Jaeyeon Jung, Stuart Schecter, Robert Biddle, Heather Lipford. Nicolas Christin, nothing brightens my day-before-a-paper-deadline more than your dry, hilarious emails. Simson Garfinkle, your advice and assertiveness is always appreciated. Diana Smetters, I appreciate that you were willing to ask me tough questions, even during lunch interviews. Mary Ellen Zurko, thanks for taking the time to understand who I am.

Golan Levin, no one could so consistently challenge me and motivate me to look towards new ideas, projects, and directions as you. So much of what I hope to be as an educator was informed by what you taught me and the courses we taught together. Most importantly though, you showed me how much can be accomplished by a strong, caring community. A community of artists, scientists, and thinkers that I think only you can assemble.

My committee:

Alessandro, your focus on the bigger picture, on what really matters, helped me take a step back from the details of whatever slight interaction or design I was testing at the time. I look forward to our continued conversations on how nudging can be used for good, not evil.

Sunny, your honesty and straightforwardness made academic work less maze-like. Working with you at Intel was, of course, an eye-opening and unique experience and I am excited for the next project that will bring us together.

Norman, you showed me how being a Computer Science professor can be about so much more than what a naive graduate student

expects, teaching me about business models and startups, patents and grants. I will be a more successful faculty member for it, and together we can continue to experiment with how to actually manage our limited time.

Lorrie, I said I wouldn't have a PhD without everyone on this list, but this holds especially true for you. Thank you for always being there, for teaching me to not write academic papers with surprise endings, supporting my endless extracurricular activities, and forcing me to concentrate and actually finish this thing.

Paul Andre, Moira Burke, Sean Munson, Jon Schleuss, Behzod Sirjani, Scott W. H. Young, you inspire me: to have more impact, to be a better educator, to find what I am deeply passionate about, to laugh more, to read more, to make connections I would never have seen.

Carrie Hagan, Warren Katzenstein, Chad Ellis, D.J. Schuldt, Hillary Schuldt, Kate Zimmerman, Patrick Zimmerman, Peter Landwehr, Amelia Kriss worked tirelessly to improve the life of graduate students through CMU's Graduate Student Assembly. Their work made me happy to serve for two years, and there is no one I would have preferred to hand the reins to than Carolyn Denomme. Through student government I was also able to meet undergraduates who juggled more than seemed possible, including Rotimi Abimbola, Aaron Gross, Sean Weinstock, Adi Jain, Nara Kasbergen, Pooja Godbole, and Jared Itkowitz.

I have to thank everyone at the National Association of Graduate and Professional Students, including Kate Allison, P.J. Dillon, Alex Evans, Kevin McComber, Denise Philpot, Christine McCary, Debra Jo Scardino, Mary Winn, Julia Mortyakova, and David Grimes. And every organization deserves a leader as organized, diligent, and hard-working as Jon Kowalski, to whom I am sure I owe a dozen favors.

The Tartan, Carnegie Mellon's newspaper, served as a second office and often a second home from the day I started graduate school. Kristen Lukiewski and Bradford Yankiver deserve all the responsibility for bringing me into the organization. But I stayed because of everyone, Cecilia Westbrook, Madelyn Glymour, Katie Chironis, Maricel Paz, Shweta Suresh, Nikunja Kolluri, Kristen Severson, Claire Morgenstern, Andrew Peters, Jessica Thurston, Emily Dobler, Anna Walsh, Nick Harper, Greg Hanneman, Alan Vangpat, Celia Ludwinski, Christa Hester, J.W. Ramp, Marshall Roch, Sarah Mogin, Alex Crichton, Josh Smith, Allison Cosby, and Isaac Jones. Courtney Wittekind and Stacey Chin not only made me show up at Sunday production, but made me stay long enough to drive them home to the apartment we shared. And as the

editing my invented words and comma-riddled sentences, Michael Kahn may deserve sainthood.

The administration and staff at Carnegie Mellon put up with my constant complaints and suggestions. I have to thank Jared Cohon, Mark Kamlet, Bob Reppe, Ralph Gross, Queenie Kravitz, Indira Nair, Renee Camerlengo, John Hannon, Kim Abel, Gina Casalegno, and Michael Murphy for letting me sit on far too many committees and put as much of the student voice into the conversation as I could manage.

The Carnegie Mellon libraries were ever-supportive of me, especially Erika Linke and Gloriana St. Clair. The library and CMU's students were lucky to have such strong leadership during a time of transition across all university libraries. And Gloriana, thank you for your insight, your advice, and always being available to chat.

The faculty and staff members at the Rochester Institute of Technology who propelled me into graduate school, taught me how to organize people, to write a passable essay and a decent poem: Babak Elahi, Linda Reinfeld, Molly McGowan, Marcia Birken, David Mathiason, Harry Schey, Larry Winnie, Cathy Winnie, Rebecca Housel, Lisa Hermsen, Katie Terezakis, and Anne Coon. Jessica Lieberman and Amit Ray continue to treat me with undeserved regard and I am ever thankful for everything they have done for me. I will visit Rochester again soon.

Tammy Stawisuck, Laurence Hiller, you two kept me occupied through high school. Norman Schoell, I never worked harder than when I was part of M.P.P. back at Ohio. High school, college, graduate school, getting a PhD, none of it was as challenging or educational as your fourth- and fifth-grade experience.

To all of my friends, thanks for being there, making me go out, play games, attend poetry readings, see movies, and enjoy life in North Tonawanda, Rochester, and Pittsburgh. Josh Debner, Arpi Kovacs, Aaron Vanderbeek, Kenny Stauffer, David Warren, Brittany Forks, Danny Rashid, Rob Simmons, Jana Diesner, Maritza Johnson, Kristen Dorsey, Stephanie Rosenthal, Dan Morris, Ben Stephens, Aaron Roth, Joseph Bradley, Jackie Libby, Greg Drozd, Kyle Orland, Drew Inglis, Corinne Walters, Eric Baldwin, Alex Loewi, Haixin Dang, Alexander Cheek, Shelly Ni, Max Hawkins, Craig Toocheck, June Bott, Cory Hoffman, Kevin Galens, Jamie Duke, Melissa Zaczek, David Riley, Greg Dufore, Kate Recard, Cheston Lee, Brian Saghy, Colin Hill, Adam Zielinski, Erhardt Graeff, Colleen Burkett, Casey LaFleur, Amber Zinni, Allison Johnston, and Anna Murray.

Katie Hempel, Phluff Kokanovich, Elise Rumpf, Carolyn Westcott, Kerri Hayes, and Amy Slevar: Thanks for being around forever.

Never go away. Remind me to visit you in DC and Hawaii and NT and the middle of nowhere. Remind me about spending summers in the pool or ridiculous driveway performances or secret forts in the basement. Remind me about the place we all came from and the memories we share. I don't ever want to forget that.

Caroline Kessler, if you keep writing me secret messages I promise to keep writing them back, someday we will have a whole volume full of the most beautiful nothings.

Elliot Norwood, keep reminding me to take breaks on the nights I don't have time to sleep and someday hit diamond.

Benjamin Foster, you helped me become a more logical, rational, emotional, better person, in a way no one else took the time to do.

Matthew Kay, you make my work more solid, methodological, and ethical. I look forward to a decade of fonts.

Daniel Freeman, if you keep trying to turn me into a computer scientist I promise to keep promising we will make something bigger.

Andy Schultz, our arguments make me so angry. Keep doing that. And keep trying to beat me at pounce.

Marissa Miracolo, I wouldn't have this PhD if we hadn't gotten ice cream together back in 2006. Our joint-insanity is almost certainly the ticket to our success.

Thank you all forever,

Patrick Gage Kelley

May 2013