“How Come I’m Allowing Strangers To Go Through My Phone?”— Smartphones and Privacy Expectations

Jennifer King
jenking@ischool.berkeley.edu

Please do not quote without permission!

ABSTRACT
This study examines the privacy expectations of smartphone users by exploring two specific dimensions to smartphone privacy: participants’ concerns with other people accessing the personal data stored on their smartphones, and applications accessing this data via platform APIs. We interviewed 24 Apple iPhone and Google Android users about their smartphone usage, using Altman’s theory of boundary regulation and Nissenbaum’s theory of contextual integrity to guide our inquiry. We found these theories provided a strong rationale for explaining participants’ privacy expectations, but there were discrepancies between their expectations, smartphone usage, and existing platform designs and data access practices by application developers. We conclude by exploring this “privacy gap” and recommending design improvements to both the platforms and applications to address it.

Author Keywords
Smartphones; touch phones; privacy; mental models, applications; apps; user expectations; iPhone; iOS; Android; Google.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Security; Human Factors; Legal Aspects.

1. INTRODUCTION
In 2012, the threshold of smartphone ownership in the U.S. surpassed 50 percent of all mobile subscribers. Google’s Android, the most popular smartphone operating system in the U.S., claimed over fifty percent of the market share, followed by iOS (32 percent).[20] Both platforms claimed over a half-million third-party applications in their respective online stores. While smartphones allowing the installation of third party applications predate the iPhone, the introduction of Apple’s iOS in 2007 ushered in a new era of application adoption by popularizing applications and attracting developers at a wide scale. Much of the smartphone’s popularity is due to its all-purpose usefulness, with users routinely using their devices to bridge diverse facets of their lives. This boundary-crossing usage leads to the broad generation and storage of personal data across multiple life dimensions.

Key elements of mobile platforms increase the risks to the privacy of the personal information users store on their smartphones. We define information privacy risk in this paper as access to one’s personal data or information without express knowledge or consent. The relative ease of developing applications for both platforms has led to an unprecedented number of developers creating them globally. Some applications need access to various categories of personal data accessible through the APIs to provide desired functionality; others do not. The amount and type of personal data made accessible to and collectible by applications by default through the platforms’ APIs also increases users’ exposure. Through the platforms’ APIs developers can access and copy user contact lists, text messages, photos, and more. Some applications need access to this data to provide functionality, but some developers access this data without a legitimate need. Both platforms seek to protect consumers’ personal data through agreements with developers that limit developers’ use of personal data obtained through the platform. But many application developers lack experience handling customer data and are unaware of the privacy and security risks and practices to manage them.

Negative public reaction to incidents where developers accessed sensitive phone data without users’ explicit consent reveals that both users and policymakers are concerned with the increased risk of collection and misuse of personal information. This also suggests that current platform designs not only do not capture users’ expectations about applications’ access to personal data, but also contradicts them. In this paper we use two theories of privacy to examine smartphone users’ concerns with other people and applications accessing the personal information stored on their smartphones. Using Irwin Altman’s theory of boundary regulation and Nissenbaum’s theory of contextual integrity to guide our inquiry, we explore the divergence between participants’ privacy expectations, understanding of risk, and management of access to their smartphones and their personal information. Using two qualitative methods—structured interviews and card sorting exercises—we explore the information privacy expectations of 24 Apple (iPhone) and Google (Android) smartphone users. Prior studies of smartphone privacy have focused primarily on concerns with location data, or have explored privacy issues within the context of device security. We cast a broader net by exploring participants’ use and
relationship to their smartphones, and participants’ privacy concerns with others’ access to varied categories of personal data entrusted to their smartphones.

2. BACKGROUND
In this section, we briefly review differences in application management between the two platforms and how they directly affect the accessibility of users’ personal information. Space constrains a detailed review, but their APIs allow access to similar information, and the data types that applications can export under both APIs’ default settings include contact lists, text messages, and photo libraries, among others.

Apple reviews each application for content, quality, and security before allowing it into their App Store (the only approved method for obtaining applications on the iPhone). Application developers must state in their terms of service and privacy policy (if they have one; many do not [30]) what data they collect and how they use it. Apple claims to reject applications with data collection policies that are “inconsistent” with the intent of the application. iOS 5 required user consent at runtime for an application to access a user’s present location (given through a dialog box), but otherwise all other data access occurred in the background. This will change in iOS 6; user consent will be required for access to additional data types, and Apple will provide an interface for reviewing or managing application access to data beyond location.[10] Prior to that, if users wanted information about the data collection and use policies of an application they had to resort to the application’s terms of service (TOS) or privacy policy if the application had one. These documents are often dozens of pages long on an iPhone’s small screen and written in legal language that is challenging for most users.

In contrast, Google does not proactively review applications for content, quality, or data collection practices, though they do scan applications for malware. Anyone can submit an application to the Google Play store (previously the Android Market), and Google also allows applications to be downloaded from external sources. During the installation process Android presents users with a non-skippable screen displaying the data categories that the application is requesting to access. The installation and consent process is binary; a user must either accept all the requested permissions or forego installing the application.

Applications running on iOS and Android routinely access and collect both personal data stored on users’ phones via their APIs as well as collect data about user behavior within applications.[29] Both the platforms and application developers have come under fire for the use of both types of data without obtaining clear consent from users. For example, in 2011 The Guardian reported that Facebook’s mobile application copied a phone’s address book to one’s Facebook account, leading to confusion and anger from users surprised by the default export of data from their phone to Facebook’s application and ultimately their online profiles.[2] Similarly, in February 2012, a software developer blogged that the social networking application Path was transmitting his entire iPhone address book to their servers without consent, noting “I feel quite violated that my address book is being held remotely on a third-party service.”[28] The observation resulted in a media firestorm. Under pressure from the U.S. Congress, Apple announced a change in practice requiring a runtime opt-in for application access to a user’s address book in iOS 6.[23]

3. RELATED WORK
Presently, there are two qualitative studies that richly document the varied contexts of smartphone use, though neither explicitly examines privacy concerns or expectations. Matthews et al.[16] and Barkhuus and Polichar[4] provide qualitative examinations of smartphone use in everyday life. Matthews found smartphone use was highly contextual, with both situation and place strongly influencing when and how people used their phones. Barkhuus focused on smartphones as ubiquitous computing devices and how “seamlessly”—or not—the devices supported users in their everyday tasks.

Häkkilä and Chatfield’s 2005 study of mobile phone privacy focuses on mobile telephone use generally and SMS messaging specifically.[13] Using both survey and interview techniques, the authors found that their subjects considered mobile phones to be “private and personal devices.” They found that privacy is protected by widely held social norms about the confidentiality of phones and messages. Until recently, most mobile privacy studies examined issues with location sharing, but a few studies focused primarily on security have also examined user privacy on mobile and smartphones. Karlson et al. in a qualitative study examined users’ preferences when sharing mobile phones with other people in order to model users’ security preferences.[14] They found that their 12 participants shared their phones with a variety of different “guest” groups, and that concerns with sharing were tied to concerns about access to the data stored on the phones. Mushukov et al. explored users’ data protection requirements on smartphones, finding that while users stored “sensitive and valuable data” on their devices, they typically did nothing to secure them from intrusion by other people.[19]

Felt, Egelman, Chin, Wagner and others have conducted a series of studies examining user perceptions, comprehension, and preferences with smartphone APIs. In [7], Felt et al. studied user comprehension and preferences with Android application permissions. The study found that only 17 percent of participants paid attention to the permission screens, and of those only 21 percent understood their content. Forty-two percent of participants were completely unaware of permissions, leading the authors to conclude, “current Android permissions warnings do not help most users make good security decisions.” Based on these findings, Felt et al. proposed a
set of guidelines in [8] to help designers in determining more effective permissions granting mechanisms in order to avoid habituation effects and alert users to potential privacy risks.

Next, Felt et al. conducted a survey asking smartphone users to rank their level of concern with 99 distinct risks associated with 54 Android API permissions. [9] They constructed a ranking of the risks, finding that users’ concerns about a specific permission and its risk were dependent upon the context of usage by an application, and somewhat surprisingly that concern about access to location information was low compared to other permissions. Finally, in [6] Chin et al. conducted qualitative interviews and surveys with 60 participants ascertaining users’ comfort levels with conducting specific tasks on smartphones to test the hypothesis that users elect not to use smartphones for some tasks due to privacy or security concerns. They also explored the reasons why users selected applications for download. Using willingness to perform the same tasks on personal laptops as a basis for comparison, they found that participants were less likely to perform privacy-sensitive tasks, such as financial and health-related transactions, on their smartphones, with over 60 percent citing security-related concerns. Participants also expressed more concern with personal privacy issues on their phones than on their laptops.

4. METHODS

Our study consisted of 24 structured qualitative interviews and a card sorting activity with 11 iOS and 13 Android users representing a range of demographic groups living in our metropolitan area. While interviews support detailed and nuanced investigation, because of the small sample size and exploratory nature of the inquiry we caution against extrapolating our findings to the broader population of smartphone users without further exploration and testing.

The interviews took place over a three-week period in August 2011. We recruited a convenience sample primarily via Craigslist, and respondents took a brief online survey to aid in screening for demographic criteria and to ensure they owned Apple or Android phones and used applications. We randomly selected respondents from pools grouped by gender, age, and phone type. We wanted to interview users of both platforms because although the platforms’ APIs provide developers with similar access to data, the manner in which users are made aware of data sharing, and the oversight practiced by the platforms themselves over applications, differ.

Respondents were told they were participating in a general smartphone study and received $40 upon completion. The interviews were approximately one hour in length, were audio recorded and later transcribed, and followed a structured interview instrument. They consisted of a series of questions about participants’ behaviors, sharing preferences, and experiences with their smartphones. Participants completed card-sorting exercises that allowed them to visually represent the way they thought about mobile applications and data sharing. We interviewed nine men and fifteen women for the study. Our participants ranged in age from 18 to 64, with nearly sixty percent under the age of 34. Education levels varied, with the majority (18) reporting completion of four years of college. The largest ethnic group (15; 63 percent) identified as White, followed by Asian-American (4), African-American (2), Hispanic (2), and mixed race (1). In a 2011 study of smartphone users in the United States, the Pew Internet & American Life Project found that in younger age brackets, higher percentages of Americans own smartphones, with the greatest between ages 18 to 29, and they are owned at greater rates in groups with higher levels of education.[27] Coincidentally, our participant pool mirrored the distribution found in the Pew study.

Interviewees had a range of work and school arrangements, and a subset of our participants worked in technology professions. None specifically worked with mobile applications or had specialized knowledge of the interview topics, though some knew developers who created mobile applications. We were sensitive to the potential bias introduced by asking participants about privacy. To avoid influencing their responses or heightening their concern, we avoided the use of the term “privacy” in our questions, phrased all descriptions about data access and sharing as factually and neutrally as possible, and placed questions that dealt the most directly with risks and concerns towards the end of the interview.

4.1 Theoretical Frameworks of Privacy

Both platforms’ notice and consent frameworks reflect the goal of facilitating individual control over the disclosure and management of personal information dominant in U.S. information privacy law. As Section 2.3 notes, the platforms’ approaches to privacy protection has not quelled the concerns of users or policymakers. We believe this is at least in part due to a failed conceptual framework—that of notice and consent. We argue that designing technical systems, including mobile systems, that afford privacy requires attention to concepts of privacy that emphasize privacy as an iterative and situated practice rather than a formalistic, procedural, mechanism. Thus our inquiry is framed around the theories of Irwin Altman—privacy as boundary regulation—and Helen Nissenbaum—privacy as contextually appropriate information flows.

Boundary regulation, as conceived by psychologist Irwin Altman to explain privacy practices in physical contexts, is “an interpersonal boundary process by which an individual or group regulates interactions with others.”[1] Individuals manage their privacy through a dynamic and contextual process of regulating personal boundaries. Leysia Palen and Paul Dourish’s framework for understanding privacy as applied to HCI is based upon Altman’s work, where they also draw on his theory to grapple with the privacy issues that arise from mediating interpersonal interactions with
information technology.[24] We used Altman’s theory to structure our exploration of our users’ decisions to allow physical access to their smartphones and potentially their personal data. Contextual integrity (CI), focuses on the contextual norms—specifically appropriateness and flow—that govern personal information.[22] We used Nissenbaum’s theory to structure our exploration of our users’ decisions to download mobile applications that implicitly or explicitly allowed access to their personal data. Louise Barkhuus recently advocated for the use of CI in HCI privacy research, and specifically studies of mobile systems, in order to better understand the “underlying contextually grounded reasons for people’s privacy concern or lack thereof.”[3]

5. FINDINGS

We begin with a brief overview of our participants’ descriptions of their relationship with and use of their phones, which documents our participants’ intense connection to and reliance on this inanimate object. We then offer a detailed account of participants’ privacy-related concerns and expectations about their phones and the data stored on them, centering on other people’s physical, and application developers’ remote, access to them.

5.1 Usage: “You panic if the phone is not with you.”

Our participants love their smartphones. As one phrased it, “I am extremely addicted to my iPhone. I use it all the time. I use it for everything, e-mail, directions, apps, reading books, listening to music. It is my life.” Smartphones are deeply integrated into their daily lives. Several expressed feeling lost without their phones: “You panic if the phone is not with you.” Twenty participants said they carried their phones at all times, and seventeen reported never turning them off: “It’s the first thing I do in the morning, it’s the last thing I check before I go to bed.” Participants used a broad range of applications and enthused about their utility and usefulness. Most reported using their phones for work and/or school in addition to their personal use, thus the device blends data and supports activities from different life domains. The boundary crossing nature of the activities they support offers one explanation for participants’ strong attachment to their phones.

Participants found and chose applications through a variety of methods, including recommendations by friends, searching the App Store or Google Play store, through companies’ websites, and top application lists from magazines, websites, and blogs. They use crowd-sourced information, such as comments/reviews and ratings, to choose applications. Some picked applications to support a particular need or desire. Participants were generally experienced with deleting applications, with the most common reason for deleting was disuse: “cleaning house.” A few mentioned privacy or security concerns as a basis for application removal.

5.1.1 Apps: Website shortcuts? Icons?

We were curious about how people conceived of applications. Functionally they are software programs, yet because users are often introduced to applications through websites, and many websites offer applications that mimic the functionality of their site, we hypothesized that some users might confuse an applications’ functionality with the browser’s. To explore this concept we asked our participants how they would describe an application to someone who had never used one before. The results were varied; only about half described an application as a software program, with more Android than iPhone users doing so. One participant explained: “I didn't understand that icons like what we have on our desktop computers and laptops, icons or quick launch, [were] different from an app. The concept of app as a mini program written for a phone didn’t sink in until a little later until I actually got the phone and I started playing extensively with downloading apps and looking at comments and seeing, oh, there [are] developers that are creating these apps.” Those who understood what applications were realized there was a difference between accessing a website through an application versus a browser, though their descriptions of those differences focused on the limited features and interface of the applications. The substantive difference noted was that an application was optimized for mobile display; no mention was made by any participants of the API as a difference between the two. The other half of our participants either confused applications with websites, describing applications as “icons on my phone” or “shortcuts to a website,” or provided answers based on applications’ functionality: “It allows you to do all sorts of things;” “It makes life a lot easier.”

5.2 Privacy Expectations

We explored participants’ expectations of privacy from two dichotomous dimensions: the risks and concerns associated with other people’s physical access to the user’s smartphone and the personal data within it, and applications gaining remote access to personal data through the API. Our theoretical frameworks guided our inquiry; we explored whether, how, and why participants regulated other individuals’ access to their smartphone (boundary regulation); and, whether expectations of appropriateness with data types and information flows (contextual integrity) informed participants’ decisions about applications. We also examined whether differences between the two platforms, such as the effect of reviewing applications prior to distribution in a platform’s store, as well as whether notice and consent mechanisms affected expectations.

We used device sharing as the entry point into our investigation of interpersonal privacy concerns because it is (generally) a voluntary, knowingly undertaken activity that nonetheless makes the sharer’s personal data vulnerable to access by another person. It offers a useful comparator case for exploring privacy concerns with the access and collection of data by application developers through the
API. For our purposes electively downloading and installing third-party applications is conceptually similar to a user’s decision about whether to allow another person to use their device as both activities pose a risk to personal data residing on the device. At the same time, we also hypothesized that the activities may be experientially distinct in ways that matter to the user’s risk perception and risk avoidance, as the consequences for exposure in each case can be substantively different.

5.2.1 Smartphone Access by Others: “It’s like going into somebody’s computer.”

Participants considered their smartphones to be highly private as well as deeply personal devices, and they articulated strong opinions about when and who they allowed to access them. “For me, it’s a very private, personal thing. There is an intimate relationship with your phone.” Another drew upon the fact that smartphones are no longer just phones: “This is not really a phone, it’s a computer. So it’s like going into somebody’s computer.”

First, we asked about voluntary sharing: whom did participants trust to share their phones? We found norms governing both with whom and when to share access. Most participants said they only shared their phones with people they knew, and expected that the person with whom they shared would only use it to make a call, look up information, or play a game—not read their emails or text messages, or look for other personal information. Participants shared their phones with family members, friends, or other trusted individuals, though a few noted that they would occasionally share a phone with a stranger or an acquaintance that needed to make a call. Some mentioned they shared their phone rarely (and unwillingly), and others claimed not to share their phone at all, even with those close to them. As one participant described it: “I don’t mind if they’re not looking at my information. If they’re just using it for the browser, [or] if it’s just the nearest phone, I don’t care at all. But if they’re getting on there to look at my emails, obviously that’s much different. But I trust my friends and family and husband enough to where I wouldn’t be worried about that.”

Next, we asked about involuntary access—someone using their phone without their permission. Nearly every participant explicitly articulated their concern in terms of access to the information on their phone, most commonly mentioning text messages, photos, email, and applications with pre-enabled logins, such as webmail or financial accounts—not the use of the phone or charges that might result. We also asked what would concern them more: a stranger accessing their phone, or someone they knew? Overwhelmingly participants were more concerned with strangers: “I’d say it’d be better off in the friend’s hands than the stranger’s hands, although personally I would keep my phone away from either.” Several volunteered that access by a stranger could only occur if their phone were lost or stolen, which gave rise to specific concerns (e.g., a stranger or thief might access one’s bank accounts, email, or social media accounts). When considering people the participants knew, several expressed qualifications regarding their level of concern, noting that it depended on the context: who the person was and why they would access their phone. For a necessary need (e.g., making an urgent phone call) they may not mind, though others said they would be upset if someone they knew used their phone without permission no matter what the circumstance. Users appear to rely on these norms to protect the privacy of information on their devices, as only two locked their phones with a passcode. Of the four who reported having had their phones stolen, none could directly trace any deleterious privacy consequences from the theft.

Overall, participants viewed their phones as private and personal closely held devices that functioned much like an extension of their selves, containing highly personal, detailed information about their lives. They relied on shared norms to protect their privacy in the limited contexts in which they shared their devices with friends and family. One participant summarized these boundary regulations concretely: “I have a lot of friends who have phones or other kinds of smartphones so we have a culture around what is permissible use because they’re at least aware of the kinds of things you can do with it and the kinds of information they store and you wouldn’t want other people to read.”

5.2.2 Data Access by Applications

We asked a series of questions to probe participants’ understanding and expectations about applications’ access to the personal data on their smartphones, exploring their privacy concerns by focusing on trust relationships and their expectations around the access, storage, and sharing of their personal data on their smartphones. We avoided direct questions about the API because they assumed technical knowledge.

5.2.3 Trusting Applications and Platforms

We asked participants whether they trusted some applications more than others. The open-ended construct of the question was chosen to allow participants to interpret the term “trust” as they wished. A minority of participants explicitly tied their perception of an application’s trustworthiness to whether it could be relied upon to access and manage their personal data fairly and respectfully: “When you engage in a relationship with this phone, with all these applications and what not, there’s an understanding here that you’re going to respect me and I’m going to respect you back.” In contrast, most participants tied trustworthiness to functional reliability, anchoring their answers in terms of how reliable they found applications (e.g., if the information an application returned was accurate, or if it was stable and didn’t crash the phone). This perception appeared to be at least partially influenced by participants’ prior experiences with the reduction in performance caused by malware on desktop systems, where
poorly functioning software was perceived as a sign of a potential system compromise. Additionally, several mentioned *name or brand recognition* as a factor influencing whether or not they found an application to be trustworthy enough to install it. Name recognition was tied to expectations around assurance structures: i.e., the "organizational or institutional mechanisms [that] exist to protect individuals from harm." [5] In this instance, the participants’ expectation was that a large company wouldn’t jeopardize a good reputation or risk a lawsuit by creating something that potentially ran afoul of the law or risked public disapproval.

5.2.4 Application Review – Taming The Wild West

We explored whether the different approaches to curating the applications available to run on the two platforms influenced user expectations of privacy. We hypothesized that review of applications by the platform owner would be perceived as an assurance structure, and therefore potentially influence users’ perceptions of privacy risks and thus application use. While Apple’s review provides some form of assurance about the applications in their store, it’s difficult for users to discern exactly what is being vouched for, since Apple is notoriously vague about the standards it uses to review and approve applications. Regardless, participant awareness of Apple’s application review was high; only two of our iPhone users were unaware of it. Strikingly, all but two of our Android users believed that Google also reviewed applications before allowing them in the Google Play store, with several mistakenly attributing Apple’s review policies to Google. A few participants mentioned that they would only install applications from the Google Play store based on their assumption that they were reviewed.

Apple’s review policy increased iPhone participants’ trust in the applications running on their phone. As one put it, “I feel like if it’s in the store, then it’s fine.” A few expressed reservations about the process, focusing on issues of content censorship by the company. When the misinformed Android users were told that Google *did not* review applications prior to their availability at the Google Play store, about half expressed a wish that Google would do so, primarily for security and privacy reasons. “I want them to review because I want them to protect me from unscrupulous data collectors. It’s a Wild West data collection issue.”

5.2.5 Disclosures – “I always click yes.”

Next, we examined the disclosures the platforms make to users. Research on online privacy policies concludes such disclosures are largely ignored.[18] We wanted to know whether smartphone users notice privacy policies and other legal disclosures related to applications, and if so, do they read them? Further, we sought to discover whether Android users notice, read, and understand the permissions presented to them during the installation process. Nearly all the participants recalled seeing a terms of service (TOS) or privacy policy in some form on their phones, but most suffered from notice fatigue: only one claimed to have read one. The majority reported clicking through or ignoring such notices, while a few reported skimming them: “I never read them but I always just click yes.” No participants reported the contents of a TOS or a privacy policy deterring installation of an application.

As discussed above, Apple requires that iOS application developers disclose data collection practices in their TOS or privacy policy (if they have one). Notices are not uniform—in language or presentation—across applications, nor do applications have a uniform process for obtaining user consent. Some applications present a runtime dialog asking users to consent to their TOS. Prior research found users to be habituated to clicking through such disclosures without reading.[11] Given the abysmal read rate of these documents reliance on them as the sole means of conveying privacy risks to users to empower them to manage their privacy risk is an unsuccessful strategy, and thus iPhone users are likely highly unaware of the extent to which data collection takes place. Nearly all the Android users recalled seeing a permission screen, but two participants reported never reading them, two thought permissions were a TOS agreement (and consequently ignored them), and two said they didn’t understand what permissions were.

The language used by Android permissions was (at the time of the study) quite technical. Only two participants felt they understood what an application was allowed to do after reading the permissions, another four reported a general understanding but stated they didn’t understand a few of them. Four participants reported, at least once, not installing an application based on the permissions it requested. Two participants noted that the language needed improvement, and one pointed out that the yes/no nature of permissions presented a non-negotiable choice that left her frustrated: “I like that I’m being asked permission but I also don’t feel like it’s really that much of a choice because you either accept it or you don’t get to use the application.” Overall, our findings with permissions complement those in [7].

5.2.6 Impressions of Data Access, Use, and Sharing by Applications

Our participants completed a card sorting exercise structured, following contextual integrity theory, to explore their preferences for the information flows from their phone to applications, application developers, and third parties with whom application developers might share or sell their information. We used a card sort to ensure that we would have a consistent method for evaluating and comparing preferences across participants. In each instance we chose two applications already installed on participants’ phones, selecting one that was account-based (requiring a login) and one that was not. We used participants’ own applications to limit friction caused by confusion about application functionality. We selected login and non-login based applications to explore potential differences between the
privacy concerns within relationships with companies that required some amount of personal information (e.g., an email address) or personalization to function, and those that did not.

We showed participants twelve cards (11 for iPhone users), one for each of the following data types: phone number, text messages, location data history, real-time location from GPS, browser cookies, browser history, photos from camera, address book, device (phone) ID, phone call logs, Apple or Google ID, and files on SD card (Android only). We explained that these cards represented different data types that were stored on their phones. We did not include the data collected individually by applications about customer usage (such as individual or aggregated usage statistics). We advised participants that if there was a data type that they were unfamiliar with to set it aside, though we provided basic definitions if asked. For each of their two applications, we conducted three card sorts (for a total of six card sorts) with each participant:

Sort 1: Which of the following types of phone data do you think this application needs in order for it to work (function) on your phone?

Sort 2: Which of the following types of data would you be comfortable with the application collecting and storing off of your phone and on the developer’s website?

Sort 3: Which of the following types of data would you be comfortable with the application developer sharing or selling with other companies?

We asked participants to select the cards with the data types they thought were relevant for each sort. Results were tallied for each card sort. Not all of the cards represent data available to applications through respective phone APIs, and between each platform there are some variations in how applications can access data. For example, we included browser cookies, location history, and Google and Apple IDs even though applications cannot access these data types because we thought participants would be familiar with them and we were interested whether participants would over or under-assume the amount of data their applications could access. We must note that at the time we conducted the study, iOS 5 was not yet available; since its release, iPhone users might have become more aware of (and perhaps less sensitive to) location requests. One reason might be the iOS runtime prompt for accessing location has made iPhone users more aware of (and perhaps less sensitive to) location requests.

Participants were more comfortable with applications receiving personal data than they were with those applications sharing it with other entities. While participants could understand the rationale for and were generally tolerant of (when it was contextually relevant) an application developer transferring data from their phones to their website, there was little tolerance of third-party data sharing—a flow of data that violated their sense of contextual integrity. During interviews, seventeen participants told us they did not want their applications to share or sell any data to third parties. One participant described his sense of violation: “I always thought these things were personal devices that we use for ourselves, for our own benefit, but apparently, people have other ideas. I guess we have to share even our own lives with these people.” Of the quarter of participants who were O.K. with Android users. Because of the overall similarity between the application types used by both groups and the generally favorable comments iPhone users had regarding Apple’s application curation policies and processes, we conjecture that iPhone users’ greater familiarity with Apple’s store review provides some rationale for these findings. Both iOS and Android users mistakenly believed that applications could access a phone’s browser cookies. They cannot, however the Android API allows access to browser history, while iOS does not. We believe the cookie misunderstanding reflects general confusion about how cookies work [17] and, as noted above, the blurry line between applications and websites for many users. Understanding some users’ lack of differentiation between accessing a website via an application, or accessing it via a mobile or desktop browser may offer a crucial insight for disentangling users’ mental models about applications’ functionality.

We tallied and analyzed the selection frequencies for each data type. Participants indicated more comfort with sharing their real-time location, device ID, and location history, and less comfort sharing their photos, address book, call logs, text messages, and files stored on SD cards (Android only). These findings were consistent with comments during the interviews tying comfort with sharing to perceptions about the level of privacy sensitivity of each of these data types, and echo the findings in [9]. Participants tied the use of data by an application to its function—demonstrating that context shaped expectations of privacy consistent with the theory of contextual integrity. Several rooted their level of comfort with an application receiving their data in the specific context of the application’s function or request, such as a banking application asking for access to one’s current location in order to locate the nearest ATM. Again, there were differences between the platforms; across all of the card sorts, for example, iPhone users selected real-time location almost twice as often as Android users. One reason the iOS runtime prompt for accessing location has made iPhone users more aware of (and perhaps less sensitive to) location requests.

Participants were more comfortable with applications receiving personal data than they were with those applications sharing it with other entities. While participants could understand the rationale for and were generally tolerant of (when it was contextually relevant) an application developer transferring data from their phones to their website, there was little tolerance of third-party data sharing—a flow of data that violated their sense of contextual integrity. During interviews, seventeen participants told us they did not want their applications to share or sell any data to third parties. One participant described his sense of violation: “I always thought these things were personal devices that we use for ourselves, for our own benefit, but apparently, people have other ideas. I guess we have to share even our own lives with these people.” Of the quarter of participants who were O.K. with
third party data sharing, location-related data was the type most frequently viewed as legitimate to share. Notably, in all instances where participants selected location data, the application was one where location sharing was relevant to the application’s function (e.g., using Yelp for local recommendations), reinforcing participants’ reliance on context to inform expectations of information sharing consistent with contextual integrity. The next most popular data type selected for sharing was device ID.¹

6. DISCUSSION

Nearly all of our participants’ expectations of privacy around access to their personal data were consistent whether the risk was from people or applications. Following Altman’s theory of boundary as well as protecting it from those who have physical access to the device. Some existing research in mobile security and privacy [14, 19] offers design suggestions for addressing privacy and security risks posed by in-person adversaries. We will focus here on a less traversed “privacy gap”—the manner in which design shapes the flow of personal data between users’ devices and applications, and the divide it creates between users’ expectations and their behavior when design fails to capture their preferences. Because participants’ sense of appropriate sharing varied based on an application’s function and context of use, the default access all applications have to personal data on both mobile platforms is a source of concern. But while platform design is crucial, it is not the only factor contributing to the gap between users’ expectations and their behavior. Researchers have documented the gap in other areas [e.g.,12], and multiple explanations have been proffered to explain it. In the smartphone context, we believe that the incongruity between preferences and behavior is composed of several competing aspects: confusion about the differences in capabilities between applications versus websites; applications’ lack of signaling when they access personal data; misplaced reliance on and false beliefs about assurance structures; and the ineffectiveness of the existing notice and consent regime. These factors contribute to an environment where users’ privacy expectations become divorced from their behavior.

6.1 Application and Website Confusion

Some of our participants did not distinguish between applications and websites. This was particularly true when using an application created by a website (e.g. Facebook.com and the Facebook mobile application); some thought an application was simply a “shortcut” to a website. Others believed that the only substantive difference between the two was the optimization of the application’s UI for the mobile platform. Even among the participants who understood that applications were not websites, most thought that applications could access their browser’s cookies, demonstrating a fuzzy demarcation between browser and application functionality. Users have been trained to look for signals that suggest privacy or data security violations, and subsequently we found that participants who assessed an application’s trustworthiness based on its functional validity referenced experiences with desktop computing systems and anti-virus software, where applications that deliver a low-quality experience or “act strange” signal a lack of credibility or potential malice. This experience is useful for evaluating applications that are overtly questionable. However, this reference point offers no help to users faced with credible, well-functioning applications that nonetheless are engaged in data access and sharing practices enabled by the API yet that violate their expectations.

If, as our research suggests, existing default accessibility of data to applications is inconsistent with users’ privacy expectations then ensuring that users understand when they are interacting with applications, versus visiting web sites, is an essential—if insufficient—step to alert them to the risk that personal data may be automatically disclosed in the background. There is currently little to no signaling to users that an application is accessing their personal data in real time. This leaves users unable to manage access to personal data in a manner consistent with their privacy needs. While presenting endless runtime prompts to users is not a desirable outcome, making application requests for personal data visible and prominent may offer promise. The challenge is to do so creatively and effectively without contributing to notice fatigue by overwhelming or desensitizing them to potential risk. But visibility without agency is not meaningful; to be effective, users would need to be given choices over the use of their data, such as the ability to deny its use by an application, rather than simply clicking a non-negotiable consent dialog more frequently.

¹It appeared few understood that the device ID could be used to track their phones by third parties; most indicated they believed it was a serial number of use only to their service provider.
6.2 Trustworthiness, Assurance Structures, and Disclosures

Most participants believed that a variety of assurance structures—e.g., obtaining applications only through official stores, or a developer’s positive reputation—protected them from privacy violations by applications. Notably, participants did not or only minimally used an application’s TOS, privacy policies, or with Android, the permission interface to evaluate the trustworthiness or privacy risks posed by applications. Though many mobile applications currently lack privacy policies [30], nearly all of our respondents reported that they usually don’t read them anyway. Further, as several of our participants pointed out, the non-negotiable, take it or leave it “consent” demanded by these tools is disempowering, especially when it is predicated upon having read and understood the aforementioned unreadable policies. This regime of false choice provides users with few meaningful options. If the only choice available is to either use an application or not, our participants generally chose to use an application and hoped that either they selected wisely or that their belief in assurance structures offered them protection.

It wouldn’t be difficult for designers to improve mobile notices by standardizing their format, making them visually comprehensible and reader friendly, though this approach may not solve the larger problem of privacy policies remaining unread.[15,25] Introducing meaningful choice for consumers by allowing them to easily and clearly set their preferences for data collection and sharing with applications may make a difference in bridging the privacy gap, which today is grounded in the “take it or leave it” approach that the Android permissions exemplify. However, improving on the implementation of the notice and consent paradigm alone is insufficient to address the privacy concerns we document; both the platforms and application developers should explicitly design for privacy concerns by drawing on the theoretical understandings of privacy we discuss here.

Users rely upon the design choices and business models of platform providers to make privacy-related decisions—and, rightly or wrongly, to protect their privacy. Although our participants used multiple features of the application marketplace—e.g., reviews, ratings—to make privacy-related decisions, these features were proxies and are not currently optimized for this purpose. But they could be; there is an opportunity for the application stores to offer tools to inform and guide users in making selection choices that support their customers’ privacy interests. Our research suggests that users expect application stores to police privacy to some extent. Given that Google Play does no curation and Apple’s curation is not aimed at protecting privacy in a manner consistent with the expectations of our participants, users’ reliance on curation to protect their privacy places them at risk. The belief that application stores provide affirmative privacy protection may contribute to user behavior that further undermines their stated privacy interests, contributing to the self-reinforcing notion that existing designs adequately address users’ privacy needs.

7. CONCLUSION

The default flows of smartphone APIs defy users’ privacy expectations. The structures in place to notify users of data flows aren’t used, and even if they were modified to be more useful they would still burden users with repeated interruptions and, ironically, potentially result in desensitizing them to the privacy risks they seek to highlight. Our participants expected that the uses of their personal data would hew closely to the minimum required for an application’s functionality. Most strongly objected to the transfer of data to third parties. These findings suggest that platform developers should restrict API defaults and design explicitly for privacy in ways that better align flows of personal data to users’ expectations. Adopting theoretically and empirically grounded design recommendations for privacy would reduce the need for policing and curation by platform providers, reduce the burden on users who don't want to trade privacy for functionality, and ease policymaker and public concern about smartphone risk. Given that smartphone use often traverses multiple personal boundaries, both individuals and institutions should consider the privacy risks posed by the default data flows enabled by platforms.

REFERENCES


