A Large Scale Study of User Behavior, Expectations and Engagement with Android Permissions

Weicheng Cao and Chunqiu Xia, *University of Toronto;* Sai Teja Peddinti, *Google;* David Lie, *University of Toronto;* Nina Taft, *Google;* Lisa M. Austin, *University of Toronto*

https://www.usenix.org/conference/usenixsecurity21/presentation/cao-weicheng

This paper is included in the Proceedings of the
30th USENIX Security Symposium.

August 11–13, 2021

978-1-939133-24-3

A Large Scale Study of User Behavior, Expectations and Engagement with Android Permissions

Weicheng Cao[*] Chunqiu Xia[*] Sai Teja Peddinti[†] David Lie[*] Nina Taft[†] Lisa M. Austin[*]

[*] *University of Toronto*
[†] *Google*

Abstract

We conduct a global study on the behaviors, expectations and engagement of 1,719 participants across 10 countries and regions towards Android application permissions. Participants were recruited using mobile advertising and used an application we designed for 30 days. Our app samples user behaviors (decisions made), rationales (via in-situ surveys), expectations, and attitudes, as well as some app provided explanations. We study the grant and deny decisions our users make, and build mixed effect logistic regression models to illustrate the many factors that influence this decision making. Among several interesting findings, we observed that users facing an unexpected permission request are more than twice as likely to deny it compared to a user who expects it, and that permission requests accompanied by an explanation have a deny rate that is roughly half the deny rate of app permission requests without explanations. These findings remain true even when controlling for other factors. To the best of our knowledge, this may be the first study of *actual* privacy behavior (not *stated* behavior) for Android apps, with users using their own devices, across multiple continents.

1 Introduction

Permission requests in the Android system have two important functions. First, they allow users to control a mobile application's ability to access resources and data on the phone. Second, they are a mechanism that informs users about the types of data that a mobile application might access. An important ramification of this system is that developers could interpret users' decisions as hints on how to develop privacy friendly applications. While many factors influence users' decisions about which permissions they grant and which they deny, this behavior could nevertheless be viewed as an opportunity to learn about unpopular permissions, which permissions make sense to users, the reasons they grant permissions and whether application-provided explanations affect users' decisions. In this paper, we focus on the permissions Android categorizes as "Dangerous", namely those which must be explicitly granted by the user to the application. Android categorizes permissions into 11 permission groups (such as *Location*, *Camera*, *Microphone*, etc.), which, for simplicity, we simply refer to as "permissions" in this paper.

Many factors affect how users interact with Android permissions, such as behaviors, expectations, explanations offered, and attitudes. Prior work usually focuses on one aspect of users at a time, such as behaviors [4, 20, 49], expectations [19, 24, 48] or attitudes [19, 35], and do not seek to analyze the interplay of these factors over the same set of users. These prior studies used surveys, or provided users with special devices, but it is preferable to obtain behavior data "in-the-wild" (when users employ their own devices) as opposed to experiments in a lab, as this captures more naturally the choices users make in their daily lives. Finally, even the largest published research studies to date that record behavior on smartphones contain at most on the order of low hundreds of participants from a single geographic region.

In order to overcome these challenges, we designed an Android app, called PrivaDroid, and used it as our study instrument. PrivaDroid is designed to run in the background on participants' phones. It observes app installs, permission grant and deny events, and launches in-situ surveys immediately after these events. Together, the observations and surveys collect data on participant decisions, rationales, expectations and attitudes at the moment they act on their own personal devices. In order to reach a broad base of participants, we designed PrivaDroid to support all major Android versions from 6.0 to 10, translated PrivaDroid into 4 major languages and used mobile advertising to recruit participants.

Our collection of decision rationales is similar to [4]; in fact, we re-use the questions from this prior study, so we can directly compare decision rationales. We expand beyond [4] in multiple ways: 1) the prior study was done with US based participants only, whereas our study includes participants from 10 countries and regions, and our app was deployed in 4 languages; 2) we collect which permissions a user expects an app to ask for and thus can compare expectations against behaviors; 3) we identify apps that provide explanations for their permission requests and those that do not, and can thus assess the impact on deny rate of providing explanations; and 4) we have users complete a privacy attitudes survey at the end of our study, so that we may compare self-stated privacy sensitivity with actual behavior.

The app was published on the Google Play Store from September 2019 to August 2020 and advertised on several online advertising platforms to recruit participants. To the best of our knowledge, this is the first cross-continent study on Android permission decision making. Over the course of our experiment, ∼1,700 participants joined from 10 countries and successfully finished the 30 day study. In total, we observed ∼72K app installs and ∼36K permission decision events. Nearly 1/3rd of these events were followed by an in-situ survey that the participant completed. This is a much larger scale study than [4] which was based on 157 participants.

Prior studies have advocated that explaining the reasons for permission requests to users is critical to improve their understanding, which in turn influences their grant and deny choices [19, 22, 28]. In previous surveys, users state they would be more comfortable granting permissions if explanations were offered [40]. Our study allows us to examine actual user behavior both in applications that offer explanations and applications that do not.

Our contributions can be summarized as follows.

- We design and implement the PrivaDroid app to collect behavioral data and perform experience sampling. We translate PrivaDroid into Spanish, French and Chinese (Traditional) and show that it is possible to use online advertising to recruit participants from around the world.

- We compare the deny rate trends today to the study done three years ago [4] and report which trends have remained the same and which have evolved.

- We find that some countries form cliques with statistically similar deny rates, but also that deny rates may differ significantly between countries in different cliques.

- Using regression modeling we show there is a statistically significant association between participants' permissions decisions (grant/deny) and their run-time expectations, as well as with their install-time expectations. We also employ these methods to show that deny rates are lower when explanations are present. These findings remains true even when controlling for other factors (such as country, attitudes, etc).

- We use a logistic regression model to study the influence of 12 factors on users' permission decision behavior. Our model shows that nearly all of these parameters have statistically significant influence on users decisions. This sheds light on the complexity of understanding user decisions as many factors play a role.

- We compare privacy attitudes to behaviors and find that ∼29% of our participants who say they are privacy sensitive also exhibit low deny rates. Analysis shows that these participants' expectations about permissions tend to be more accurate (matching app behavior), suggesting that privacy sensitive users who grant many permissions may be doing so with a better understanding of how and why applications use permissions.

The rest of the paper is organized as follows. Section 2 discusses related work. Section 3 explains the participant recruitment method, while Section 4 describes the design, data collection and implementation of our PrivaDroid app. Our findings are presented in Section 5. Section 6 describes the limitations of our study and Section 7 concludes the paper. The survey questions are listed in Appendix A.

2 Related Work

There is an extensive amount of existing research in the space of Android permissions and privacy. Much of this research documents user discomfort with permissions [35, 49] and their frustration with what appears to them as unnecessary permission requests [9, 20, 44, 46]. This can happen because developers are not knowledgeable about permissions and this results in mistakes [35, 38], or (mis)use of permissions in unexpected ways [29, 39]. A recent study has shown that developers mostly use default configurations when integrating ad/analytic libraries, and choose these libraries based on popularity and ease-of-use [26]. Many studies have found cases where app permission requests are not related to the app's core functionality [1, 6, 19, 29, 31, 32, 34, 36, 46]. We do not focus on developers in this work, but instead on users.

Research on user privacy expectations with permissions has shown that users are concerned when they learn of the possible risks associated with permissions [11], or about applications collecting data when running in the background [13, 42]. In [19], the authors studied user expectations around 4 resources (GPS location, Device ID, network location, contact list) based on an older model of Android. This study captured resource requests users did not expect via an mTurk survey, not based on decisions on personal devices as in our study. Wijesekera et al. [48] captured user expectations by monitoring their apps for one week and showing users afterwards what was collected and asking in-lab questions about whether the participants expected that. This study reports that users said they were more likely to deny permissions they didn't expect. Our results corroborate this finding, however we use a very different mechanism as we captured actual decisions made on personal devices, and at a much larger scale.

To help provide explanations or additional information so users can make better choices, Harbach et al. [12] and Kelley et al. [15] suggested providing more privacy information and personal examples to help improve user comprehension.

Others categorized permissions to reduce the number of privacy/security decisions users need to make [10]. Some have explored creating personalized privacy assistants [20], or surfacing nudges to assist users with decision making [2]. This research focuses on supplementary features to help users make decisions, whereas we focus on developer provided explanations.

There is little work on app-provided permission explanations. Tan et al. [40] conducted an online survey of smartphone users and showed that permission requests that include explanations are significantly more likely to be granted. They also analyzed ∼4K iOS apps and showed that only 19% of the permission requests included text within the dialogs to explain the request. Liu et al. [21] analyzed ∼83K Android apps and the extracted permission explanation messages, and showed that less than 25% of apps provide explanations and that the purposes stated in a significant proportion of these explanations were incorrect. We have made similar observations in our analysis too: only 15% of apps in our data presented an explanation to users for their permission requests, and having an explanation reduced the permission deny rate from 15.4% to 7.1%. While the prior work mentioned influence of permission explanations on the denial rates based on surveys, ours is the first to study user behavior on their own devices and quantify the reduction in actual permission denial rates when explanations are present.

Others have conducted cross-country studies [3, 5, 7, 14, 25, 27, 30, 33, 35] related to privacy. For example, Shklovski et al. [35] conducted interviews and a survey across two countries (Iceland and Denmark) to investigate how smartphone users feel about data access on their phones and if they are willing to change their behavior after being informed about tracking and data leakage. A multi-country survey [25] showed that psychographics and various attributes of the mobile app context are predictive of users' privacy preferences. Schubauer et al. [33] examined app behavior on the Google Play Store across three categories and 3 countries (US, South Korea and Germany) and discovered that policy changes aligned with privacy law changes (such as the General Data Protection Regulation) have impact on application permission usage. Overall, there has been little research comparing users in different countries in terms of their attitudes and behaviors related to Android app permissions. With the exception of [33] that focuses on app design, the other prior studies use interviews and surveys as their methodology. To the best of our knowledge, we are the first to compare *actual* privacy behavior (not *stated* behavior), with users employing their own devices, across multiple countries.

3 Participant Recruitment

Participant Composition. We recruited participants from 10 countries and territories, namely Canada, United States, Argentina, United Kingdom, France, Spain, South Africa, India, Singapore, and Hong Kong. These countries were selected using multiple criteria. First, we aimed to cover a diverse set of regions thus selecting countries from 5 continents, covering 4 languages, and with different privacy legislation. Second, we selected countries where we had access to native speakers of the dominant language spoken, enabling us to check our translations. Third, we focused on countries with high smartphone penetration [37] and included two developing economies, South Africa and India. Finally, we aimed to include countries covering a range of privacy views: India previously had low privacy awareness and few concerns about privacy [17, 18] whereas France and Spain are reputed to have strong concerns about privacy and are in a region (Europe) with some of the strictest privacy laws (GDPR). This ensemble of countries is similar to that in [5] which also includes 2 or 3 countries each from Europe, North America, Asia and 1 from South America.

Our aim was to recruit at least 100 participants from each region with a nearly balanced split between males and females, hoping to obtain sufficient data to compute statistically significant results. Because participants self-enrolled asynchronously, and advertisements are sent out in large batches, we could not control the number and gender of participants who joined our study, and this created variance in participant numbers across countries. We found that females were less likely to join our study despite efforts to target more advertisements at females. We did not control for other variables, such as age, profession or income during the recruitment process, mainly due to the inaccuracy in the advertisement network inferred attributes for targeting our ads and partly due to ethical concerns over targeting for age or income.

Advertising and Compensation. We use online advertising to recruit participants as it allows us to find participants across many countries. Most recruitment agencies for user studies only work in a single country, and international ones are prohibitively expensive—particularly for large studies. We selected three popular online advertising providers, namely Google, Facebook and Reddit, so as to reach a broad audience. Initial experimentation with our app revealed that male participants were more likely to join our experiment than females. To improve gender balance we targeted our advertising towards female participants first, and only started advertising to males after we had more than 50 female participants.

We offered participants $10 USD if they stayed for 30 days and completed the experiment. We initially selected Bitcoin and PayPal as payment methods. However, Bitcoin was not approved by our IRB, so we used PayPal for all participants.

Transparency and User Consent. *This study was approved by our institutional review board (IRB).* Participants need to give their consent before enrolling in our experiment. This process happens after they install and open the PrivaDroid Android app. The consent form enabled us to both gain consent and allowed us to be transparent about our practices. It contains the following key clauses. First, participants must come from one of the specified countries and must be above 18 years of age. Second, participants must keep the accessibility service and app usage access enabled for our app during the length of the experiment. Finally, we notify participants that PrivaDroid collects no personally identifiable information except for their Google advertising identifier (a device ID that we use to associate all the data coming from a single device), and that we don't use this advertising identifier to infer any other personal information (such as name, email, etc). Participants must consent to these clauses.

Data Protection. To protect user privacy, access to the collected raw data is controlled, and limited to only the subset of authors (at the University of Toronto) directly involved in the implementation and maintenance of PrivaDroid.

4 PrivaDroid Data Collection Platform

The PrivaDroid data collection platform consists of an Android application and a backend that stores and analyzes data. PrivaDroid is designed to collect both behavioral data on certain events and in-situ survey responses right after those events occur. PrivaDroid manages and tracks participant participation over the course of the study. We describe the data we collect, how we design surveys and how we localize PrivaDroid into 3 other languages. Technical details of PrivaDroid can be found in Appendix B.

4.1 Behavioral Data

The PrivaDroid application records participants' app install and permission decision events, as well as the permission rationale dialogs shown by the app. For app install events PrivaDroid logs the app's package name or the application ID, its version info and the title. PrivaDroid records app updates from app installs separately, but this study only considers app installs and ignores updates. PrivaDroid logs permission decision events that happen in the runtime permission request prompts, as well as decisions that occur when the user navigates to the Android Settings Menu and toggles an app permission there. For each permission decision event, PrivaDroid captures the aforementioned app information, the permission type being requested by the app or being modified by the participant, as well as whether the participant granted or denied the permission.

Some apps use a custom dialog that provides an explanation for a permission request along with a set of buttons for the users to indicate whether they are willing to grant the permission. If the user agrees, the app will subsequently request the permission via Android system. However, if the user does not agree, the app will not request the permission. This has the side effect of reducing the number of permission requests made by an application via the system APIs, and causing under-counting of the number of permission denies, since PrivaDroid's monitoring of permission decisions via the system APIs doesn't capture deny events occurring indirectly in custom dialogs. To measure this effect, as well as measure the frequency of applications using such permission explanation dialogs, PrivaDroid captures the text on these dialogs using a keyword-based heuristic and the accompanying button that was clicked. We evaluate the accuracy of our heuristic in Section 5.3.1.

4.2 Survey Design

Participants answer three types of surveys in the PrivaDroid app (provided in Appendix A). First, PrivaDroid uses a survey to capture the demographic information of our recruited participants. Participants are required to take this survey right after sign-up. They are asked to provide their age, gender, income and education level. We use this data to analyze and compare behavior and privacy perspectives across different demographic groups.

Second, PrivaDroid presents in-situ surveys that are designed to capture either the participant expectation, comfort, or decision rationale at the moment a relevant event occurs. At app install time, we invoke one survey (Appendix A.2) to capture expectations about permissions before the app is used. Other surveys are invoked right after a permission grant or deny event, so we can capture participant rationales, runtime expectations, comfort, and desire to grant temporarily (Appendix A.3 and A.4). Following best practices, we impose a limit of a minimum of 5 minutes between consecutive in-situ surveys to avoid overloading participants [45].

Last, participants are required to answer an exit survey at the end of the 30 day study to complete the experiment and receive the compensation. The survey derives questions from the well established IUIPC privacy scale [23]. The questions are used to compute a *privacy score* for each participant along the four dimensions: Control, Awareness, Collection and Secondary Use. As the IUIPC scale was originally developed in 2004 and focused on general "Internet use", we adapted the questions in a minor way to focus on mobile privacy. Specifically, we replaced the term "online companies" with "smartphone apps", and replaced the term "consumer online privacy" with "mobile app privacy". Our 15 questions (See Appendix A.5) were scored on a 5-point Likert scale, as opposed to the original 7-point scale as we learned that multilingual surveys are more frequently done with 5-point scales [47]. We mapped the answers to the range $\{-2,2\}$. To evaluate the quality of our mobile-specific IUIPC questions, we conducted a 100 person Amazon Mechanical Turk survey and the resulting Cronbach's Alpha scores in the range of 0.65 to 0.82 demonstrate acceptable reliability. Both the PrivaDroid and mTurk surveys include a simple attention check question to ensure that participants are actually reading the questions, and we discard the data of participants who fail to correctly answer the question.

4.3 App Localization

In order to include non-English speaking participants, we translated and localised PrivaDroid into Chinese (Traditional), Spanish and French. The translation consists of two parts: 1) strings in the PrivaDroid app, such as the consent form, the survey questions and answers, etc.; and 2) strings in the Android System UI, such as those used in detecting the permission changes participants made on the Android Settings page, Android system runtime permission dialogs and participants' decisions. For the first part, we used the translation service provided by the Google Play Console and then had native speakers check the translations. For strings involved in the Android System UI, we used the translations provided in the open-sourced Android framework Git repositories.

5 Findings

5.1 Data Summary

We advertised our study on the three advertising networks mentioned earlier, across 10 countries and regions, and ran it from Nov 2019 to May 2020. As mentioned before, we initially targeted our ads towards females to encourage their participation. After onboarding 50 or more females per country, we relaxed the targeting criteria and showed ads to all. Hong Kong was the only region where we did not reach 50 female participants; thus we use the Hong Kong data for aggregate analysis hereafter, but not for demographic analysis. In total we spent $12,953.85 USD on advertising to recruit participants, which generated 2,640,029 impressions, 20,947 clicks and 5,377 installs of the PrivaDroid app. Of the installs, 1,719 participants stayed for the required 30 day period to complete the study. 1,044 of our participants identified themselves as males, 655 as females, and the rest identified as neither or preferred not to state their gender. Another 2,207 participants joined the study but withdrew before 30 days, thus we exclude their data. (Many participants installed the app but didn't join the study.) Table 1 summarizes some participant demographics. During the study period, these participants carried out 72,214 app install events of which 36% were surveyed, and 36,152 permission decision events of which 30% were surveyed. Due to our self-enforced limitation on how frequently surveys were shown to participants each day, not all events result in a survey being triggered.

5.2 Permission Denials

Of the ∼36K permission decision events across the 11 permission groups, we found that our participants denied 16.7% of these permission requests. Even without considering the events for recently introduced permissions (such as *Body Sensors*, *Physical Activity* and *Call Logs*), the average deny rate is very close to the 16% reported in an earlier study [4]. In our current study, we observed that 8% of the permission decisions occurred from the Settings menu, which is similar to the 5% reported in [4]. For these two aggregate metrics, the behavior has not changed much since 2017. Of all the decisions our participants made via the Settings menu, 40% were to deny a previously granted permission. While this number is high, it still means that the majority of decisions made at the Settings menu were to grant a permission. As we will see shortly, a top reason for denying a permission is because participants are aware that they can go to the Settings menu and change their decision afterwards.

Country and Region	Males	Females	Other	Prefer not to say
Canada	107	75	5	1
US	99	132	3	3
Argentina	175	57	0	1
UK	86	57	0	0
France	97	53	1	0
Spain	126	82	1	3
South Africa	56	70	0	0
India	187	57	0	0
Singapore	59	52	0	1
Hong Kong	52	20	0	1
Total	1,044	655	10	10

Table 1: Country and Gender Demographics

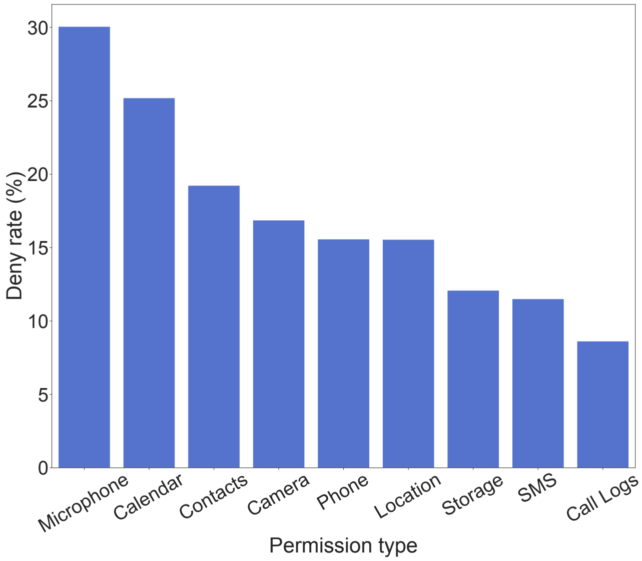

Figure 1: Permission deny rates for each permission group

Both the number of events and deny rates vary a lot based on the individual permission type. *Storage*, *Location*, and *Camera* were heavily requested with each having >5K events. However, we saw very few permission decision events for *Body Sensors*, *Call Logs* and *Physical Activity* permissions - perhaps because these three permissions were fairly new (at the time of our study).

Figure 1 shows deny rates for each permission group (we only include those with at least 50 decision events, thus eliminating *Body Sensors* and *Physical Activity*). *Microphone*, *Calendar* and *Contacts* have the highest deny rates of 30%, 25% and 19%, respectively. Permissions such as *Location* and *Storage*, which are the most frequently requested in our data, have lower deny rates of 15% and 12%. The average deny rate across all permission requests was 16.7%. Compared to deny rates recorded in [4] (which only included US participants), we see that deny rates for our US participants have increased for *Calendar* (to 21.7% from 10%) and *SMS* (to 15.6% from 10%), and decreased for *Phone* (12.6% from 19%), *Location* (8.5% from 15%), and *Camera* (11% from 15%).

About 11% of our participants had Android 10 devices, giving them access to the *foreground only* permission option introduced in it. Although deny rates for the *Location* permission on Android 10 and earlier were roughly the same at 17% and 15%, two thirds of the *Location* permission grants in Android 10 were *foreground only*. This suggests that users not only want to be able to control if location can be used, but when it is used as well. Since the option is only available for *Location* permission and in Android 10 alone, which did not make up a big portion of the collected data, we treat *foreground only* option as a permission grant in this paper.

In examining the rationales our participants gave for denying permissions, we see that the top three reasons for denies are: "I can always grant it afterwards if I change my mind" (27% of denies), "I do not use the specific feature associated with the permission" (25% of denies), and "I think the app shouldn't need this permission" (23% of denies). The first reason indicates that participants are aware that they have the ability to revise their permission grant and deny decisions, while the second and third reason demonstrate that users may be trying to enforce the principle of least privilege either based on their usage of the application or their understanding of the operation of the application. These rationales illustrate that users think about app functionality and app features they use, when permission requests are made; this kind of thinking relates to expectations that we analyze in Section 5.3.

Our top reasons are the same as those found in [4] (see Table 5 therein) with minor shifts in frequency. We test the null hypothesis that the reasons for participants denying permissions in both our experiment and in [4] are from the same distribution using a Two-Sample Kolmogorov-Smirnov (K-S) test (using the data in Table 5 of [4]). The resulting Kolmogorov–Smirnov statistic (D value) is 0.375 with a p-value of 0.66. We thus accept the null hypothesis that the distribution of deny rationales has essentially remained the same as in [4], and the top reasons remain unchanged.

Similarly, the top reasons for permission grants include: "I want to use a feature that needs this permission" (37% of grants), "I think the app won't work otherwise" (25% of grants), and "I trust the developer" (23% of grants). These top reasons are the same as those indicated in [4]. Trust in the developer still seems to play an important role in whether participants decide to grant a permission to an app. To compare the histograms of grant reasons across the two studies, we form a null hypothesis that the frequencies at which grant reasons were selected in both [4] and our experiment are from the same distribution. We again conducted a two sample K-S test and obtained a D value of 0.125 with a p-value of 1. Since the p-value is larger than the critical value of 0.05, we cannot reject the null hypothesis, and thus conclude that the frequency of how often each grant reason was chosen in our experiment is consistent with [4].

Overall, the top reasons for both grants and denies suggest that participants tended to rationalize their permission granting and denying as a trade-off between functionality and privacy. Reasons that suggest a more emotional response, such as "I have nothing to hide" or "I wanted the permission screen to go away" were chosen less often.

Temporary permissions. We also asked participants each time after they granted a permission, if they would have liked to grant it temporarily. We found that 24% of the times participants chose to grant a permission, they would have preferred to do so temporarily. Among the permissions that were surveyed at least 50 times, the desire to grant temporarily ranged from 21% to 26% depending upon the permission. In line with this, the Android 11 OS release [41] includes a *one-time* grant option for *Location*, *Microphone* and *Camera* permissions.

One could interpret the desire to grant temporarily as a hesitation, or lack of comfort, in granting a permission permanently. To check this, we first compared how comfortable participants felt when granting permissions with their desire to grant them temporarily. In the cases when participants indicated they were not interested in granting a permission temporarily, 53% of them selected that they felt either very or somewhat comfortable granting those permissions. However, among those who said they would have liked to grant the permission temporarily, only 36% of them felt very or somewhat comfortable. To determine the influence of user comfort level on the desire to grant temporarily, we carried out mixed effects logistic regression on the grant surveys. In the mixed effects logistic regression, we treat the participants' indicated comfort level on the 5-point Likert scale as numeric independent feature in the range $[-2,2]$ and the desire to temporarily grant as the dependent feature. We include the permission type as a fixed effect to control the influence of different number of events for each permission, and the participant and app as random effects so that the latent individual differences between participants and apps are taken into account in the form of different intercepts for each participant and app. The trained model shows a significant difference due to comfort (β = -0.429, p-value = $< 2e-16$). An ANOVA test between this model and a base model differing in only the comfort feature has shown that including the comfort feature did lead to a better model fit (p-value = $< 2.2e-16$). These results indicate that users' desire to grant a permission temporarily is higher when they are more uncomfortable.

5.3 Explanations and Expectations

Intuitively, the context in which a permission request is made should have an effect on whether the average user will grant or deny the request. Here, we define context as the *explanation* (if any) given at the time of the request, as well as any *background information* the application has imparted to the participant leading up to the request. While PrivaDroid captures explanations at the time of the request, background information is beyond what PrivaDroid can possibly capture, as it includes all previous interactions the participant had with the app, as well as auxiliary information such as documentation on the application's Google Play Store page, third-party reviews of the app, or even informal recommendations through friends. Nevertheless, to ignore background information would be perilous, as we feel that background information may have a strong effect on a user's disposition towards a permission request, and may even compensate for a weak or complete lack of an explanation at the time of the request. Thus, as a proxy for background information, we collect via surveys, the participant's *expectation* of a permission request at two points in the user's interaction with the application.

The first point where PrivaDroid measures expectation is during app install, when participants are asked "which of the following permissions do you think the app requires?" and they select as many as they want from the full list of permission groups. (See Appendix question A.2.1.) The second point is after the participant has responded to a runtime permission prompt, when they are asked "did you expect the app to request this permission?" (regardless of whether the participant granted or denied the permission). For this question, participants select either "Yes" or "No". While expectation cannot explain how a participant received their context (i.e. how they came to expect or not expect a permission request), these two measures approximate the participant's context from installation time up to the point that the permission request is made. Together with the presence of an explanation taken at the time of the request, we have three measures of the context a participant experiences for a permission request.

5.3.1 Explanations

As mentioned in Section 4.1, PrivaDroid collects data on permission explanation messages in the form of text dialogs shown by the app with some UI elements (such as buttons). PrivaDroid captures these explanations by scanning for Android TextViews that occur right before a permission request, and capturing those that contain a verb that is related to data collection and a noun that belongs to a permission. We then associate this explanation message with the respective permission request. We also record the button options present on the dialogs and what was clicked by the study participant (to determine if the participant approved/denied the request).

Because the collection technique relies on heuristics, it may miss some explanations. To measure the accuracy of our heuristic we perform offline analysis across 15 popular apps on the Android playstore. We run the app with PrivaDroid installed and record the screen. We then playback the recording and identify all possible explanations provided by the popular app and compare it against the captured explanations by PrivaDroid. In total we identified 30 explanations across the 15 popular applications with 22 of those captured by our heuristic. We note that we only encountered one false positive (collected by PrivaDroid but is not actually an explanation). This experiment shows that our heuristic is a conservative detector and our collected data underestimates the number of permission requests with explanations.

In total, we collected 1804 permission explanation messages that preceded a grant or a deny across 1097 apps. Thus, 15% of apps in our study include an explanation for at least one of their permission requests. It is difficult to measure the quality of an explanation from just the dialog text, as this misses any images that may be in the dialog, as well as the overall context in which the dialog is shown. We thus only examine the correlation between deny rates and the *presence* of an explanation and find that having an explanation reduced the permission deny rates to 7.1% as compared to the 15.4% deny rate for requests with no explanations. To determine if the presence of an explanation affects participants' decision to grant or deny a permission request, we carried out mixed effects logistic regression analysis due to the presence of multiple observations from each participant and for each app. We treat the presence of explanation as a binary independent feature and the permission decision ('1' represents a deny and '0' a grant) as the dependent feature. Similar to the case of temporary permission grants earlier, we include the permission type as a fixed effect and the participant and app as random effects. The trained model shows a significant difference between the presence and absence of an explanation (β = -0.854, p-value = $< 2e-16$). An ANOVA test between this model and a base model differing in only the explanation feature has shown that including the explanation feature did lead to a better model fit (p-value = $< 2.2e-16$). These results and the negative coefficient indicate that the presence of explanation reduces the deny rate for a permission request.

Explanation message dialog may cause a runtime permission request to be omitted. For instance, an app might indicate that it would like to "Use Location to show personalized ads?" with two buttons: "Not Now" and "Yes". Clicking on "Not Now" conveys to the app that the user is going to deny the permission request, so the app may simply skip making the request. Because PrivaDroid computes deny rates based on Android system permission requests, PrivaDroid will undercount these app-specific permission deny events. To adjust for this, each of the 2643 English explanation messages where

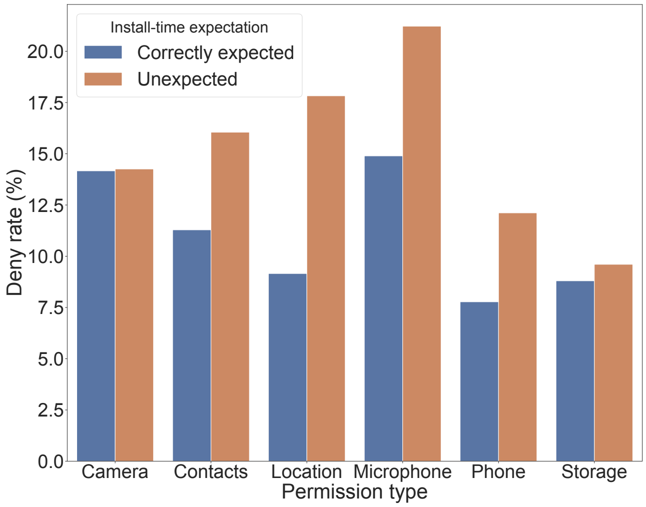

Figure 2: Permission deny rates for install-time expectations

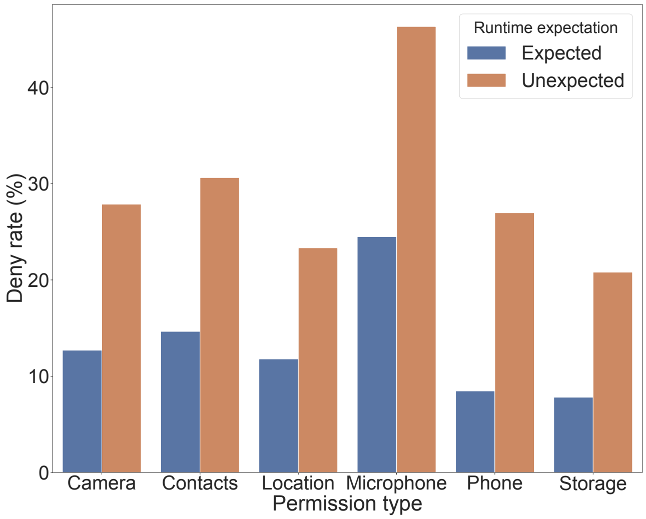

Figure 3: Permission deny rates for run-time expectations

a "Not Now" or an equivalent option was selected by the participant was manually evaluated by two of the authors to determine if it is indeed a permission rationale message, resulting in 540 actual pro-active deny messages[1]. Because this behavior only affected 15% of applications seen in our study, we use unadjusted deny rates in the remainder of the paper.

5.3.2 Context Through Expectations

Install-Time Expectations. An app may not need to provide an explanation if the user has enough context at the time of the permission request. To approximate this context, we measure user expectations of permission requests. We use the term *correctly expected* for cases when the participant expected a particular permission would be requested and the app requested it, the term *incorrectly expected* for cases when a participant expected a permission but the app did not request it, and the term *unexpected* for cases when a participant did not expect the permission, but the app actually requested it.

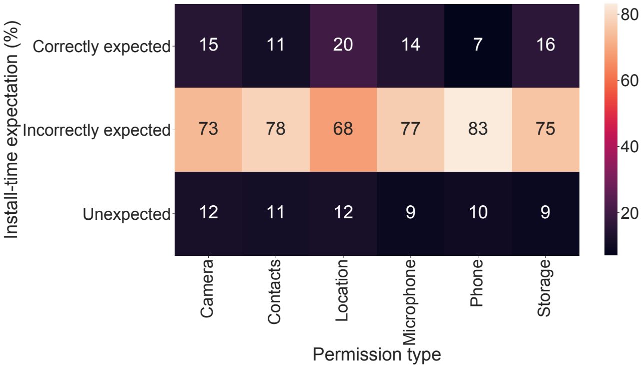

Figure 4: Permission expectations vs reality

We first examine whether our participants' install-time expectations match reality. Figure 4 shows rates for the three types of expectations for the 6 permission types with the most permission request events. The rate at which participants correctly expect future permission requests ranges from 7% for the *Phone* permission to 20% for the *Location* permission; these results suggest that at install time, participants do not have enough context to give them an accurate picture of an app's permission needs. We hypothesize that this behavior might come from participants becoming habituated to assuming that apps frequently request unnecessary permissions [9, 20, 44, 46]. Overall, this suggests that users do not have the context necessary to expect permission requests before an app is installed.

Figure 2 shows the deny rate for correctly expected and unexpected permission requests for individual permissions. (Note we cannot compute deny rate for incorrectly expected permissions since the app doesn't ask for a permission in this case.) Deny rates are always higher for unexpected permission requests, which agrees with previous research [48]. The average deny rate for expected permissions is 10.2%, whereas the average deny rate for unexpected permissions is 14.2%. This phenomenon of participants denying unexpected permissions more frequently holds in aggregate and across permission types. In order to see if the participants' install-time expectations affect their permission decisions, we again carried out a mixed effects logistic regression analysis. Not all participants shared their permission expectations at install time, so we modeled install-time expectation as a categorical feature with three levels – *Yes, No* and *Not Surveyed*; and *Yes* is chosen as the reference level. We modeled install-time expectation as the independent feature and the permission decision as the dependent feature. Similar to the earlier analysis, we include the permission type as a fixed effect and the participant and app as random effects. The trained model shows a significant difference between expecting and not expecting

[1] Some of the explanations were actually permission requests by web pages in a browser

Country and Region	Avg # of Grants	Avg # of Denys	Avg Deny Rate	Intra-country Deny Rate Std Deviation	Avg Privacy Sensitivity
Canada	15.22	3.55	18.9%	20.5%	1.25
US	27.21	3.72	12.0%	12.6%	1.10
Argentina	9.77	3.25	25.0%	25.2%	1.19
UK	16.30	3.09	15.9%	19.8%	1.13
France	12.37	2.85	18.7%	17.3%	1.00
Spain	13.10	4.14	24.0%	21.0%	1.16
South Africa	16.07	2.60	13.9%	14.1%	1.39
India	31.58	4.86	13.3%	14.7%	1.16
Singapore	13.69	2.58	15.9%	22.7%	1.29
Hong Kong	6.28	3.05	32.7%	30.0%	1.18
Overall	17.51	3.52	16.7%	6.1%[2]	1.17
Gender	Avg # of Grants	Avg # of Denys	Avg Deny Rate	# of Participants	Avg Privacy Sensitivity
Male	18.41	3.48	15.9%	1,044	1.13
Female	15.99	3.51	18.0%	655	1.25
Other	23.40	4.00	14.6%	10	1.30
Did not say	13.56	5.78	29.9%	10	0.84
Education level	Avg # of Grants	Avg # of Denys	Avg Deny Rate	# of Participants	Avg Privacy Sensitivity
Less than high school	14.49	2.46	14.5%	146	1.07
High school	17.86	3.19	15.2%	945	1.18
Bachelor's or more	17.29	4.14	19.3%	555	1.20
Did not say	20.36	5.01	19.8%	73	1.06

Table 2: Permission Request Events and Decisions

a permission at install-time ($\beta = 0.37$, p-value = 0.000451 for *No* categorical response). An ANOVA test between this model and a base model differing in only the install-time expectation categorical feature has shown that including the install-time expectation did lead to a better model fit (p-value $= 5.931e-11$). These results and the positive coefficient indicate that a permission is more likely to be denied when it is unexpected at install time.

Runtime Expectations. In 7,711 (72%) of our surveyed runtime permission events, participants expected the permission request and in the remaining 28% they did not. The number of permission events where an initially unexpected install-time permission request changed to an expected request at runtime (over all permission events where we recorded both install-time and runtime expectations) was 25% (1,233/4,892) demonstrating that users sometimes revise their expectations as a result of additional context acquired through use and interaction with an app. The deny rate for permissions expected at runtime was 12.2% whereas the deny rate for runtime unexpected permission requests was 26.9%. This $\sim$15% difference in deny rates is 3× larger than the $\sim$4% discrepancy observed for install-time expectations—participants are 2× more likely to deny permission requests they did not expect at runtime than at install-time. Figure 3 shows that the deny rate for unexpected permission requests is roughly double that of expected requests, *across all the permission types*. In the case of the Phone permission, the deny rate tripled, going from 9% to 27%. The ensemble of these observations shows that expectations do influence participant behavior, and also suggests that better understanding and more accurate expectations gained through context cause users to grant permissions.

Similar to our assessment of the influence of install-time expectations, we check if the participants' run-time expectations affect their permission decisions via a mixed effects logistic regression analysis. We again modeled run-time expectation as a categorical feature with three levels – *Yes, No* and *Not Surveyed*; and used *Yes* as the reference level. We modeled run-time expectation as the independent feature and the permission decision (recall, '1' represents a deny) as the dependent feature. Similar to the earlier analysis, we include the permission type as a fixed effect and the participant and app as random effects. The trained model shows a significant difference between expecting and not expecting a permission at run-time ($\beta = 1.21$, p-value = $< 2e-16$ for *No* categorical response). An ANOVA test between this model and a base model differing in only the run-time expectation categorical feature has shown that including the run-time expectation did lead to a better model fit (p-value = $< 2.2e-16$). These results indicate that an unexpected permission at run time makes it more likely to be denied. Our findings corroborate the findings in [48], although as pointed out in Section 2, our study mechanisms are quite different and our study size here is two orders of magnitude larger.

5.4 Cross Country Analysis

We now look at behaviors according to country and regional differences. We acknowledge that understanding country to country comparisons is challenging as it is not possible to control for all factors influencing such comparisons. Cultural values [3, 7, 27] and regulatory frameworks [33] are considered macro-environmental factors that have been shown to influence users' privacy concerns and their behavior in response to data requests. One aspect of culture, namely individualism versus collectivism, has been demonstrated [30] to influence self-disclosure. Views towards government [8] also influence privacy attitudes. A study of 7 European countries [27] showed how local culture influences privacy attitudes and stated behavior, while [7] made similar observations for large cities in 4 Asian countries. Studying cultural issues is complex in part because privacy attitudes are evolving worldwide [14]. For example, [14] reports that differences across 25 countries, in terms of how important privacy is, are minor. However, views about how privacy will improve over the next decade are significantly different across countries. While all of these factors may influence participant behavior, we could only control for the gender of our participants, and thus exogenous factors, such as Android phone popularity, and the economic value of $10 within a country, may bias the set of participants in our survey. While we may refer to the participants by the country they are from, we acknowledge - as a result of the above limitations - that we can only make observations about the participants in our study, and that disambiguating the effect of a country's culture from the other mentioned factors is beyond the scope of this paper.

Our cross country comparison includes 9 countries (recall that we leave Hong Kong out here since we were unable to recruit at least 50 female participants). Table 2 shows the deny rates across different countries, as well as (for completeness), gender and education. The aggregate deny rate per country varies from 12% for the United States to 25% for Argentina. It is noteworthy that some regions (Argentina and Spain) have deny rates that were twice as high as other regions (the US and India). However this aggregate deny rate may hide variation among participants within countries.

We perform country pairwise ANOVA tests to determine if the participants from two countries are drawn from populations with the same mean deny rates. After doing this for all pairs of countries, we identified 2 distinct cliques of countries; for all pairs within the same clique, the null hypothesis holds, indicating that the countries within a clique are similar with respect to their means. However for all pairs of countries from different cliques, the null hypothesis is rejected indicating that their populations have different mean deny rates. The US, India and South Africa formed one clique and these 3 countries have an average deny rate of 13.07%. Canada, Argentina, Spain and France belong to the second group with an average deny rate of 21.65%. Singapore and the UK did not fit cleanly into either clique. For example, although the UK was statistically similar to both France and South Africa, it differed from both the US and Spain.

Figure 5 presents the deny rates for individual permissions by country. We see that the permission type that a population is most sensitive too (highest deny rate) varies across countries. For example, *Microphone* is the most frequently denied permission in 5 countries, *Calendar* is the top denied permission in 3 countries, and *Location* has the highest deny rate only in Spain. Within individual countries, we see certain permissions are more vigorously denied than others (e.g. the French deny *Calendar* twice as often as *Camera*).

5.5 Factors Influencing Deny Rate

In Section 5.3, we used mixed effects logistic regression to study the influence of a single factor on the permission decision. In this section, we now build a larger model, that helps determine the influence of each of the dozen factors collected in the study while controlling for other factors. Similar to the earlier exercises, we consider permission decision as the binary response variable ('1' represents a deny and '0' an accept), and include the participant and app as random effects.

We consider the following factors. Each participant in our study was required to answer an exit survey that measured their privacy attitudes along the 4 dimensions of Control, Awareness, Collection and Secondary use of private information, as described in Section 4.2. Based on their responses to these questions, participants are assigned a score on a scale between $[-2, 2]$ in each dimension, with positive scores indicating higher sensitivity to privacy loss in that dimension. We included these four privacy dimensional scores (*control*, *awareness*, *collection* and *secondary_use*) as quantitative variables. The presence/absence of a permission explanation string (*has_explanation*) and the permission change happening from the settings menu (*settings_menu*) are included as binary variables. The rest of the 6 variables are included as categorical variables with reference levels. The reference levels were selected randomly to prevent any bias: "US" for *country*, "Bachelorś degree (e.g. BA, BS) or higher" for *education*, "Male" for *gender*, "Below 30" years for *age*, "Location" for *permission*, and "Yes" for *runtime_expected*. We include all the users who answered demographic questions and their permission decision events in this analysis, and not just the surveyed ones. For the unsurveyed decisions, the *runtime_expected* variable is specified as 'Not Surveyed'. Some of the categorical levels for age and education have been merged to account for low response volumes, and rows corresponding to 'Other' and 'Prefer not to say' in the gender category have been excluded from the analysis.

We performed Variance-Inflation Factors (VIF) analysis to check for multicollinearity among the 12 chosen variables. VIF measures how much the variance of any one of the coefficients is inflated due to multicollinearity in the overall model. VIF values above 5 are considered problematic. All of our 12 variables have VIF values below 5. In fact, almost all of the variables have values close to 1, except for the four privacy dimensional scores which have scores close to 4. Overall this indicates that participant demographics, their privacy attitudes, expectations, country as well as explanations and permission

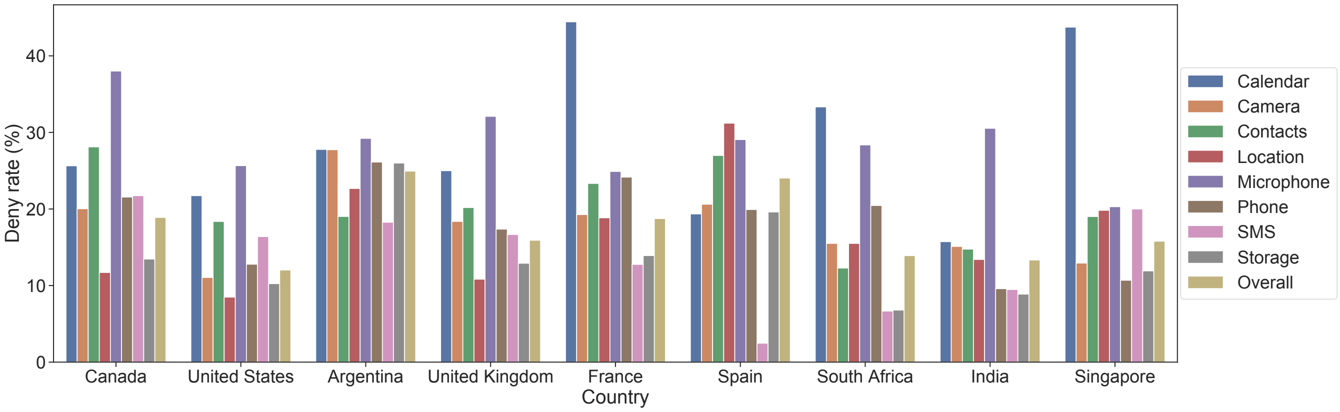

Figure 5: Permission deny rates of individual permission types in each country

types all play a role in permission denial decisions.

The results of the mixed effect logistic regression analysis with all the 12 variables and the random effects is shown in Table 3. Each row contains a factor, its accepted values, the identified β coefficient value indicating directional change in the permission deny rates with respect to the baseline of the given factor, and the p-value indicating statistical significance. Many of the factors have statistical significance with p-values < 0.001. The model has a conditional R^2 value of 0.576. The intraclass correlation coefficient (ICC) for the user random effect is 0.256 and for the app random effect is 0.271, indicating that the permission decisions from a particular user or app are not strongly correlated with other decisions from the same user or app. The table corroborates a number of our earlier findings. In Table 2 we reported higher average deny rates for women than men. With our current larger regression model, we see that females are more likely to deny a permission (β = 0.299) compared to the reference male category, when controlling for other variables. Section 5.3.1 indicated that the presence of an explanation reduces the deny rate. Our larger regression model again shows that providing a permission explanation string makes it less likely to deny the request (β = -0.725) when compared to the case where there is no explanation. Section 5.3.1 showed statistical significance between runtime expectations and the denial rate. The current larger model again shows that an unexpected runtime permission is more likely to be denied (β = 1.216), even when controlling for other factors. These results strengthen our earlier findings, as they remain true even when controlling for other variables.

Table 3 also provides additional insights. Controlling for other variables, a permission change happening from the settings menu is more likely to denied (β = 2.04). Looking at the privacy scores, users with higher sensitivity across collection (β = 0.404) dimension are more likely to deny requests, and those with higher sensitivity across secondary use (β = -0.264) are less likely to deny. When we look at the influence of a country in our data, compared to a user in the US, those coming from Argentina, Canada, Spain, France, UK and Singapore are more likely to deny a permission. India and South Africa don't exhibit statistical significant difference compared to the reference country US, perhaps because they are both in the same clique (see Section 5.4). We tested other models using different references countries (e.g. Argentina, France) and in those models, India does exhibit statistically significant different behavior. This shows that country plays an important role in permission decisions.

Users with less than high school diploma education level are less likely to deny permissions compared to those with a Bachelor's or higher degree. This finding indicates that education level does have an influence on a user's permission choices. When comparing across different permission types, our model shows that Android users' behavior does vary by permission. We see that Contacts and Microphone are generally denied more often than Location—even when controlling for a multiplicity of factors. Overall participants deny Storage less often than any other permission.

We explored whether permissions are treated differently in different countries by training a second mixed effect logistic regression model of permission deny rates with the 'country:permission' interaction effect. An ANOVA test between this model and the earlier model without the interaction term shows that the second model has better fit (p-value = 8.8e-13). This demonstrates there is an interplay between how different permissions are perceived across countries.

From this second model, we observe that some country:permission interaction variables diverge significantly (p-value < 0.05) from overall country patterns. For example, our Spanish participants generally deny permissions more compared to those in the US, yet they deny individual permissions such as Camera, Contacts, Microphone, and Storage less compared to the US. Similarly our Argentinian participants deny more than their US counterparts, but have lower denial rates for Contacts and Microphone. In conclusion, it is interesting to note that there are not just a couple of factors that influence a user's permission decision, and the final observed decision is a combined effect of many factors put together.

Variable	Values	β Coefficient (p-value)
control	[-2, 2]	-0.044
awareness	[-2, 2]	0.109
collection	[-2, 2]	0.404 (***)
secondary_use	[-2, 2]	-0.264 (*)
has_explanation	Binary	-0.725 (***)
settings_menu	Binary	2.04 (***)
country/region (reference: US)	Canada	0.870 (***)
	Argentina	0.555 (**)
	UK	0.567 (**)
	France	0.795 (***)
	Spain	0.883 (***)
	South Africa	0.068
	India	0.118
	Singapore	0.42 (.)
age (reference: Below 30 years)	Between 30 and 50	-0.104
	Above 50	-0.006
education (reference: Bachelor's degree or higher)	Less than a high school diploma	-0.249 (*)
	High school degree or equivalent	-0.193
gender (reference: Male)	Female	0.299 (**)
permission (reference: Location)	Calendar	0.259
	Camera	0.011
	Contacts	0.258 (**)
	Microphone	0.606 (***)
	Phone	-0.093
	SMS	-0.265
	Storage	-0.379 (***)
runtime_expected (reference: Yes)	No	1.216 (***)
	NotSurveyed	0.306 (***)

Significance codes: $p < 0.001$ (***), $p < 0.01$ (**), $p < 0.05$ (*), $p < 0.1$ (.)

Random Effect	Variance
App (Intercept)	1.889
User (Intercept)	1.785

Table 3: Regression Analysis to Predict a Permission Deny

5.6 Engaged Users

As described in Section 5.5, we score each participant on a scale between $[-2,2]$ along the 4 dimensions of Control, Awareness, Collection and Secondary Use based on their exit survey responses. We average out these dimensional scores, and assign an *overall privacy score* to each participant. This overall privacy score summarizes the privacy sensitivity of the user. The participants who failed the attention check question were not included in the privacy score computation.

To understand the relationship between participants' privacy scores and their permission deny behavior, we plot the distribution of the 1,027 participants who had over 10 permission events by their deny rate and overall privacy score in Figure 6. The color density indicates the number of participants in each cell. From this, we make three observations. First, as expected, as overall privacy sensitivity increases, so does the average deny rate, with an increasing number of participants having a deny rate greater than the mean (16.7%). Second, the variance of permission denying behavior increases as overall permission sensitivity increases, with high variability of deny rates for participants with high sensitivity. Finally, and most interestingly, the distribution of deny rates for participants with relatively high overall privacy sensitivity (> 1.0) is not uniform—a large proportion, 29% (296/1027), have deny rates lower than the population average of 16.7% and show up as a concentration of users near the bottom middle right.

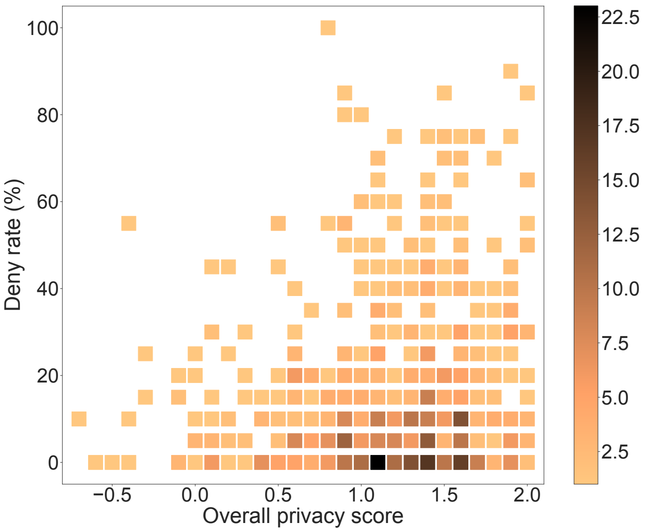

Figure 6: Participant Distribution: Deny Rate & Privacy Score

This discrepancy between the high privacy scores (attitudes) and the low deny rates (behavior) might hint at the well known "privacy paradox" effect [16, 43]. However, the permission deny/accept decisions are very contextual [28] and it is impossible to make an assessment of the privacy paradox effect without knowledge of the complete context that led to a permission deny/accept decision.

Users with high privacy scores may still allow permissions if they select their apps carefully and have a good understanding of permissions and their purpose. We hypothesize that among participants with high privacy scores, there may be users that are more *engaged*, in that they are careful in their app selection and understand context better. These engaged users might be making context specific permission choices. While another study would be needed to evaluate this hypothesis, we check if our data can offer any preliminary insights.

As described in Section 5.3.2, the participants' *expectation* serves as a proxy for the context of a permission request (where the context includes a user component in addition to the information provided by the app). We evaluate if the install time and run time expectation distributions vary between participants with high (> 1.0) and low privacy scores by performing two separate K-S tests, one for each expectation. In both the tests, we consider the null hypothesis to be that the distributions are same across the two privacy score groups. For the install time expectation case, K-S statistic (D) is 0.104 with a p-value of 0.02. For the runtime expection, D value is 0.12 with a p-value of 0.014. Based on these p-values, we can safely reject the null hypotheses and conclude that both the expectation distributions are statistically different for low and high privacy score participants. We observe that participants with high privacy scores (> 1.0) generally have higher percentage of *correctly expected* permissions at install time (average is 31.4%, median is 26%) compared to those with low privacy scores (average is 27.1%, median is 18%). Also, participants with higher privacy scores on average report higher percentages (75%) of expected permissions at runtime compared to those with low privacy scores (69%). In summary, we find that for participants with higher privacy scores their (install and run-time) expectations are more likely to match reality, than for participants with lower privacy scores. These findings partially support our hypothesis that users who have both high privacy scores and low deny rates may be more engaged as they appear to understand context better.

6 Limitations

Due to the nature of our participant recruitment, which relies on online advertising, our study is biased towards users who interact with online ads. Naturally, all of our participants were also willing to install an application that collects data about their smartphone usage. This introduces unavoidable selection bias that is inherent to our methodology, as we are unable to collect data from potential participants who do not interact with online ads or who were unwilling to install PrivaDroid. As mentioned in Section 3, we also find that females were under-represented with our methodology. To get a sense of demographic sample bias, we compared the distribution of ages and educational attainments of our US participants with US Census Bureau statistics from 2019[3]. We found that younger people (77% of study participants are under 40 vs 40% for all US residents) and those with lower educational attainment (78% of study participants have highschool or less vs 54% for all US residents) are over-represented in our group. We speculate that this bias may be due to the higher rate of smartphone use among the younger population, as well as the low monetary incentive ($10 USD) being more attractive to participants with lower educational attainment.

PrivaDroid cannot collect data on events that occurred before it was installed, thus we do not see any permission decisions participants made with their apps before the start of our study. It may thus under-count events caused by the default apps that come with a phone, or popular applications that are likely to have been already installed on a participant's phone. Both participant bias and blindness to pre-install events are unavoidable side-effects of our recruitment and data collection methodologies. In addition, 42% of the users participating in our study did so after March 15, 2020, when the social and economic measures caused by the Covid-19 pandemic came into force in the majority of the countries in our study, and we are unable to conclusively ascertain the effect of those measures on this group of participants.

From our experience with PrivaDroid, we believe a mobile application-based data collection platform coupled with advertising is a viable method for conducting global user-studies. However, one challenge we think could be better addressed in future work is techniques to more holistically collect and measure a user's context when interacting with apps. PrivaDroid focused mainly on the text in pop-up dialog boxes, but misses other important factors, such as images and general text in UI screens that are not in dialog boxes. In addition, while 36K permission request events may seem like a lot of data, it is a tiny number compared to the large variety of smartphone apps available. As a result, we have very little data on any specific app, making contextual analysis of behavior across apps impossible. For example, the largest number of permission events with an explanation for a single app in our dataset is only 54, and the number falls off fairly steeply. To better understand contextual behavior, either more data needs to be collected or the study has to be re-designed to focus on participants who use a specific subset of apps.

7 Conclusions

We have found that a few trends reported in [4] remain the same three years later: the aggregate denial rate still hovers around 16-17%, *Microphone* is still the most often denied permission, and we continue to see variation in deny rates across the permission types. At the same time, there were some notable changes for specific permission types. For example, the deny rate for the *Calendar* permission has grown significantly from 10% [4] to 21.7% today and the deny rates for the *Phone* permission have dropped from 19% to 12.6%.

Our demographic analysis reveals interesting trends across countries. We found two distinct cliques of countries in our data, where countries within a clique have statistically similar deny rates. Some countries do not fit cleanly into either clique. We also observed different permission sensitivities across countries. Previous studies [3, 7, 27] show that nationality influences users' willingness to share their personal data. Our regression models corroborate this specifically for Android apps and for user behavior on their personal devices. Our study revealed that users are less likely to deny permission requests when explanations are present. We demonstrated this trend with regression models that show this holds, even when accounting for all other factors influencing decisions (such as age, app, country, attitude, etc). The average deny rate was reduced by half when there is an explanation (15% vs 7%). Our study also shows that expectations have a significant influence on permission decision making. We found that participants deny permissions more often when an app asks for a permission they did not expect. We again demonstrated this via

[3] Statistics from https://census.gov/data/tables/2019/demo/age-and-sex/2019-age-sex-composition.html and https://www.census.gov/data/tables/2019/demo/educational-attainment/cps-detailed-tables.html

regression modeling. This bias exists for both types of expectations (install-time and run-time) and across all permission types, but is significantly stronger for runtime expectations, where the deny rate for unexpected permissions is double that of expected permissions. This corroborates prior work [48] but on a larger scale and across multiple countries.

One of the main forward-looking take-aways from our study is that users are more likely to grant permission requests that are expected. In a sense, this tells us that the permission system is working—when a permission request "makes sense", users are more likely to grant the permission. This further suggests that the gap between smartphone user's desires to constrain applications and the reality is more due to shortcomings in their understanding of the interplay between apps and permissions, and the context in which permission requests are made, than the permission mechanism itself (with the exception of temporary permissions, which our study showed have some benefit to users). As a result, transparency features, such as Apple's "Privacy Nutrition Labels" and Google Play's Safety directive [4], may serve to complement the current smartphone permissions system design. However, our results also show that the effect of unexpected permissions at run-time is more pronounced than at install-time, suggesting that transparency features that only target install-time permissions may not be as effective as those that are more dynamic and linked to specific permission types. Future research on the quality of explanations and exactly how and when to use them would be very beneficial to the proper adaption of explanation labels.

Acknowledgements

We acknowledge the feedback of the anonymous reviewers and our shepherd, Sven Bugiel, whose insights and suggestions improved the paper greatly. We would also like to thank Nicolas Papernot, Adelin Travers, Beom Heyn Kim, Sukwon Oh, and Mariana D'Angelo for helping us check our translations. We thank Kelly Hayward and Dubravka Burin, who helped us navigate the numerous administrative and reporting complexities of making hundreds of international payments from a public institution. We also thank Manya Sleeper and Micha Segeritz for helping us with the regression analysis. Finally, financial support for this research was provided in part by the University of Toronto's Connaught Fund, a Canada Research Chair, an NSERC USRA grant and by a Security and Privacy Research Award from Google.

References

[1] Y. Agarwal and M. Hall. ProtectMyPrivacy: detecting and mitigating privacy leaks on iOS devices using crowdsourcing. In *Proceedings of MobiSys*, 2013.

[2] Hazim Almuhimedi, Florian Schaub, Norman Sadeh, Idris Adjerid, Alessandro Acquisti, Joshua Gluck, Lorrie Faith Cranor, and Yuvraj Agarwal. Your location has been shared 5,398 times!: A field study on mobile app privacy nudging. In *Proceedings of CHI*, 2015.

[3] Steven Bellman, Eric J. Johnson, Stephen J. Kobrin, and Gerald L. Lohse. International differences in information privacy concerns: A global survey of consumers. *The Information Society*, 20, 2004.

[4] Bram Bonné, Sai Teja Peddinti, Igor Bilogrevic, and Nina Taft. Exploring decision making with Android's runtime permission dialogs using in-context surveys. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 195–210, July 2017.

[5] Surveillance Studies Centre. The Globalization of Personal Data (GPD) Project International Survey on Privacy and Surveillance, 2013.

[6] Saksham Chitkara, Nishad Gothoskar, Suhas Harish, Jason I. Hong, and Yuvraj Agarwal. Does this app really need my location?: Context-aware privacy management for smartphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3), September 2017.

[7] Hichang Cho, Milagros Rivera-Sánchez, and Sun Sun Lim. A multinational study on online privacy: global concerns and local responses. *New Media & Society*, 11, 2009.

[8] Rowena Cullen. Citizens' concerns about the privacy of personal information held by government: A comparative study, Japan and New Zealand. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, 2008.

[9] Adrienne Porter Felt, Erika Chin, Steve Hanna, Dawn Song, and David Wagner. Android permissions demystified. In *Proceedings of the 18th ACM Conference on Computer and Communications Security*, 2011.

[10] Adrienne Porter Felt, Serge Egelman, Matthew Finifter, Devdatta Akhawe, and David Wagner. How to ask for permission. In *Proceedings of 7th Usenix conference on Hot Topics in Security (HotSec)*, 2012.

[11] 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 *Proceedings of the Second ACM Workshop on Security and Privacy in Smartphones and Mobile Devices*, 2012.

[12] Marian Harbach, Markus Hettig, Susanne Weber, and Matthew Smith. Using personal examples to improve risk communication for security & privacy decisions. In

[4] https://android-developers.googleblog.com/2021/05/new-safety-section-in-google-play-will.html

The 32nd Annual ACM Conference on Human Factors in Computing Systems, 2014.

[13] Jaeyeon Jung, Seungyeop Han, and David Wetherall. Short paper: Enhancing mobile application permissions with runtime feedback and constraints. In *The Second ACM Workshop on Security and Privacy in Smartphones and Mobile Devices*, 2012.

[14] Patrick Kelley. Privacy, measurably, isn't dead. In *Usenix Enigma*, February 2021.

[15] Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. Privacy as part of the app decision-making process. In *The SIGCHI Conference on Human Factors in Computing Systems*, 2013.

[16] Spyros Kokolakis. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & security*, 64:122–134, 2017.

[17] Ponnurangam Kumaraguru and Lorrie Cranor. Privacy in India: Attitudes and awareness. In *In The 2005 Workshop on Privacy Enhancing Technologies*, 2005.

[18] Ponnurangam Kumaraguru and Niharika Sachdeva. Privacy in India: Attitudes and Awareness v2.0. Technical report, Precog-TR-12-001, Precog@IIIT-Delhi, 2012. http://precog.iiitd.edu.in/research/privacyindia/.

[19] Jialiu Lin, Shahriyar Amini, Jason I. Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. Expectation and purpose: Understanding users' mental models of mobile app privacy through crowdsourcing. In *The 2012 ACM Conference on Ubiquitous Computing*, 2012.

[20] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun (Aerin) Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *The 12th Symposium on Usable Privacy and Security(SOUPS)*, 2016.

[21] X. Liu, Y. Leng, W. Yang, W. Wang, C. Zhai, and T. Xie. A large-scale empirical study on Android runtime-permission rationale messages. In *2018 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pages 137–146, 2018.

[22] Xueqing Liu, Yue Leng, Wei Yang, Chengxiang Zhai, and Tao Xie. Mining Android app descriptions for permission requirements recommendation. In *IEEE International Requirements Engineering Conference*, pages 147–158, 08 2018.

[23] Naresh K. Malhotra, Sung S. Kim, and James Agarwal. Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale and a Causal Model. *Information Systems Research*, December 2004.

[24] Nathan Malkin, Julia Bernd, Martiza Johnson, and Serge Egelman. What can't data be used for? Privacy expectations about smart TVs in the USA. In *European Workshop on Usable Security (EuroSEC)*, 2018.

[25] Andrew McNamara, Akash Verma, Jon Stallings, and Jessica Staddon. Predicting mobile app privacy preferences with psychographics. In *The 2016 ACM on Workshop on Privacy in the Electronic Society*, page 47–58, 2016.

[26] Abraham H. Mhaidli, Yixin Zou, and Florian Schaub. "We can't live without them!" app developers' adoption of ad networks and their considerations of consumer risks. In *Fifteenth Symposium on Usable Privacy and Security*, August 2019.

[27] Caroline Lancelot Miltgen and Dominique Peyrat-Guillard. Cultural and generational influences on privacy concerns: a qualitative study in seven European countries. *European Journal of Information Systems*, 2014.

[28] Helen Nissenbaum. Privacy as contextual integrity. *Washington Law Review*, 2004.

[29] Elleen Pan, Jingjing Ren, Martina Lindorfer, Christo Wilson, and David R. Choffnes. Panoptispy: Characterizing audio and video exfiltration from Android applications. *Proceedings of Privacy Enhancing Technologies Symposium*, 2018.

[30] Clay Posey, Paul Benjamin Lowry, Tom L Roberts, and T Selwyn Ellis. Proposing the online community self-disclosure model: the case of working professionals in france and the U.K. who use online communities. *European Journal of Information Systems*, 19, 2010.

[31] Zhengyang Qu, Vaibhav Rastogi, Xinyi Zhang, Yan Chen, Tiantian Zhu, and Zhong Chen. Autocog: Measuring the description-to-permission fidelity in Android applications. *The ACM Conference on Computer and Communications Security (CCS)*, November 2014.

[32] Jingjing Ren, Martina Lindorfer, Daniel J Dubois, Ashwin Rao, David Choffnes, and Narseo Vallina-Rodriguez. Bug fixes, improvements, and privacy leaks. a longitudinal study of PII leaks across Android app versions. In *Network and Distributed System Security Symposium (NDSS)*, 2018.

[33] Jonathan Schubauer, David Argast, and L Jean Camp. Lessig was right: Influences on Android permissions. In *TPRC48: Research Conference on Communications, Information and Internet Policy*, 2018.

[34] Fuming Shih, Ilaria Liccardi, and Daniel Weitzner. Privacy tipping points in smartphones privacy preferences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, page 807–816, 2015.

[35] Irina Shklovski, Scott D. Mainwaring, Halla Hrund Skúladóttir, and Höskuldur Borgthorsson. Leakiness and creepiness in app space: Perceptions of privacy and mobile app use. In *The 32nd Annual ACM Conference on Human Factors in Computing Systems*, 2014.

[36] Anastasia Shuba, Evita bakopoulou, and Athina Markopoulou. Privacy leak classification on mobile devies. In *Workshop on Signal Processing Advances in Wireless Communication (SPAWC)*, 2018.

[37] Laura Silver. Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center, February 2019.

[38] Matthew Smith. Usable Security – The Source Awakens. *Usenix Enigma*, 2016.

[39] R. Stevens, J. Ganz, V. Filkov, P. Devanbu, and H. Chen. Asking for (and about) permissions used by Android apps. In *2013 10th Working Conference on Mining Software Repositories (MSR)*, 2013.

[40] Joshua Tan, Khanh Nguyen, Michael Theodorides, Heidi Negrón-Arroyo, Christopher Thompson, Serge Egelman, and David Wagner. The effect of developer-specified explanations for permission requests on smartphone user behavior. In *The SIGCHI Conference on Human Factors in Computing Systems*, page 91–100, 2014.

[41] Permission updates in Android 11: One-time permissions. https://developer.android.com/preview/privacy/permissions, 2020. Accessed: June 2020.

[42] Christopher Thompson, Maritza Johnson, Serge Egelman, David Wagner, and Jennifer King. When it's better to ask forgiveness than get permission: Attribution mechanisms for smartphone resources. In *The Ninth Symposium on Usable Privacy and Security*, 2013.

[43] Janice Y Tsai, Serge Egelman, Lorrie Cranor, and Alessandro Acquisti. The effect of online privacy information on purchasing behavior: An experimental study. *Information systems research*, 22(2):254–268, 2011.

[44] Lynn Tsai, Primal Wijesekera, Joel Reardon, Irwin Reyes, Serge Egelman, David Wagner, Nathan Good, and Jung-Wei Chen. Turtle guard: Helping Android users apply contextual privacy preferences. In *Symposium on Usable Privacy and Security (SOUPS)*, 2017.

[45] Niels van Berkel, Jorge Goncalves, Lauri Lovén, Denzil Ferreira, Simo Hosio, and Vassilis Kostakos. Effect of experience sampling schedules on response rate and recall accuracy of objective self-reports. *International Journal of Human-Computer Studies*, 125:118–128, 2019.

[46] Timothy Vidas, Nicolas Christin, and Lorrie Cranor. Curbing android permission creep. In *The Web*, volume 2, pages 91–96, 2011.

[47] Ana Villar. *Agreement answer scale design for multilingual surveys: Effects of translation-related changes in verbal labels on response styles and response distributions*. PhD thesis, University of Nebraska, 2009.

[48] Primal Wijesekera, Arjun Baokar, Ashkan Hosseini, Serge Egelman, David Wagner, and Konstantin Beznosov. Android permissions remystified: A field study on contextual integrity, 2015.

[49] Primal Wijesekera, Arjun Baokar, Lynn Tsai, Joel Reardon, Serge Egelman, David Wagner, and Konstantin Beznosov. The feasibility of dynamically granted permissions: Aligning mobile privacy with user preferences. In *The 38th IEEE Symposium on Security and Privacy*, 2017.

A Survey Questions

Here we list the English version of the survey questions and available options.

A.1 Demographic Survey

Users were required to answer all questions but were allowed to select the "Prefer not to say" option.

A.1.1 What is your age?

- Below 20
- Between 20 and 30
- Between 30 and 40
- Between 40 and 50
- Between 50 and 60
- Above 60
- Prefer not to say

A.1.2 What is your gender?

- Male
- Female
- Other
- Prefer not to say

A.1.3 Which country do you live in?

List of all countries.

A.1.4 What is the highest degree or level of school you have completed?

- Less than a high school diploma
- High school degree or equivalent
- Bachelor’s degree (e.g. BA, BS)
- Master’s degree (e.g. MA, MS)
- Doctorate (e.g. PhD)
- Prefer not to say

A.2 Expectation Survey at App Install Time

We ask users right after they install an app about which permissions they expect the app will ask for. Participants can choose as many as they like.

A.2.1 Which of the following permission do you think the app requires?

- Camera
- Contacts
- Location
- Microphone
- Phone
- Storage
- Body Sensors
- Calendar
- SMS
- Call Logs
- Physical Activity
- None
- I don’t know

A.3 Permission Grant Event Survey

We randomized the order of the possible options except for the "None" and "Other", which were always placed at the end.

A.3.1 Why did you grant the permission request?

- I want to use a feature that needs this permission
- I trust the developer
- I think the app won’t work otherwise
- I have nothing to hide
- The developer already has this information about me
- I want the permission screen to go away
- Because the app is popular
- The app gave an explanation that made sense
- None
- Other

The following question is used to gauge permission expectations at runtime.

A.3.2 Did you expect the app requests this permission?

- Yes
- No

A.3.3 How comfortable do you feel granting this permission request?

- Very uncomfortable
- Somewhat uncomfortable
- Neutral
- Somewhat comfortable
- Very comfortable

A.3.4 Do you want to grant the permission temporarily?

- Yes
- No

A.4 Permission Denial Event Survey

We randomized the order of the possible options except for the "None" and "Other", which were always placed at the end.

A.4.1 Why did you deny the permission request?

- I think the app shouldn’t need this permission
- I can always grant it afterwards if I change my mind
- I do not use the specific feature associated with the permission
- I consider the permission to be very sensitive
- I don’t trust the developer
- I wanted the permission screen to go away
- The app gave a poor explanation
- I think something bad might happen
- None
- Other

A.4.2 Did you expect the app requests this permission?

- Yes
- No

A.5 Exit Survey

Users were asked to state how much the agree or disagree with each statement in this survey using the following 5 options:

- Strongly Agree
- Agree
- Neither Agree Nor Disagree
- Disagree
- Strongly Disagree

Note that question 4 in the Control section (A.5.1) is the opposite of the statement in question 4 of the Collection Section (A.5.3). This was inserted as an attention checking question. Surveys with contradictory answers were not used.

A.5.1 Control Section Questions

1. Mobile app privacy is about a user's right to exercise control over decisions about how their information is collected, used, and shared.
2. User control of personal information is essential to mobile app privacy.
3. I believe that mobile app privacy is compromised when the user loses control over their information as a result of app usage.
4. I'm not concerned that smartphone apps are collecting too much personal information about me [5].

A.5.2 Awareness of Privacy Practices Section Questions

1. Mobile app developers seeking information should disclose the way the data are collected, processed, and used.
2. A good mobile app privacy policy should have a clear and conspicuous disclosure.
3. It is very important to me that I am aware and knowledgeable about how my personal information will be used.

A.5.3 Collection Section Questions

1. It usually bothers me when smartphone apps ask me for personal information.
2. When mobile apps ask me for personal information, I sometimes think twice before providing it.
3. It bothers me to give personal information to so many mobile apps.
4. I'm concerned that smartphone apps are collecting too much personal information about me.

A.5.4 Secondary Use Section Questions

1. Mobile apps should not use personal information for any purpose unless it has been authorized by the individuals who provided information.
2. When people give personal information to a mobile app for some reason, the app developer should never use the information for any other reason.
3. Mobile app developers should never sell the personal information in their computer databases to other companies.
4. Mobile app developers should never share personal information with other companies unless it has been authorized by the individual who provided the information.

B PrivaDroid Technical Details

PrivaDroid supports Android versions starting from Android 6.0, in which the runtime permission system was introduced, up to Android 10. The data is stored off device in a Firebase cloud datastore. When one of the app install or permission decision events happen, PrivaDroid will detect it and create a notification that can direct the users to the corresponding survey. App install events are detected by listening to the `ACTION_PACKAGE_ADDED` Android system broadcast intents. For devices running Android 8.0 or higher, we additionally implemented a foreground service to listen to these broadcast events. Package name, app name and the app version name are logged by probing the Android Package Installer.

Capturing permission events is a bit more challenging as no system intent is broadcast when a permission request is granted or denied. A seemingly obvious way to observe permission changes is to consistently poll the permissions granted for an app using the Android `getInstalledPackages` API and check if anything changes. However, this approach can only capture permission changes but not the permission decisions that do not result in any change. For example, denying a request for a permission that was already denied before will not be caught. Instead, PrivaDroid uses the accessibility service facility to monitor the screen that participants view and looks for UI elements with specific strings or View IDs to detect permission prompts, and uses the app usage permission to detect which app package requested the permission. Based on the information, it then extracts the app name, permission name and the participant's grant/deny decisions.

We only use data of participants who have the application continually installed for the length of the study. However, participants may leave the study anytime they wish and they do not need to explicitly inform us when they leave. To detect if a participant has left the study, PrivaDroid implements a heartbeat message that will be sent to its Firebase datastore daily. The heartbeat message contains two booleans, which are whether the accessibility service and app usage access are enabled for PrivaDroid. These two booleans were used to determine if a user's data is valid and should be included in our analysis. Additionally, if PrivaDroid detects that either the accessibility service or app usage permission has been revoked, but the PrivaDroid app has not been uninstalled, the PrivaDroid app will create a notification prompting the participant to re-enable those capabilities in case they were disabled by accident.

[5] This is the attention-checking question.