Permission vs. App Limiters: Profiling Smartphone Users to Understand Differing Strategies for Mobile Privacy Management

Ashwaq Alsoubai
atalsoubai@knights.ucf.edu
University of Central Florida
Orlando, Florida, USA

Reza Ghaiumy Anaraky
rghaium@clemson.edu
Clemson University
Clemson, South Carolina, USA

Yao Li
yao.li@ucf.edu
University of Central Florida
Orlando, Florida, USA

Xinru Page
xinru@cs.byu.edu
Brigham Young University
Provo, Utah, USA

Bart P. Knijnenburg
bartk@clemson.edu
Clemson University
Clemson, South Carolina, USA

Pamela J. Wisniewski
pamela.Wisniewski@ucf.edu
University of Central Florida
Orlando, Florida, USA

ABSTRACT

We conducted a user study with 380 Android users, profiling them according to two key privacy behaviors: the number of apps installed and the Dangerous permissions granted to those apps. We identified four unique privacy profiles: 1) Privacy Balancers (49.74% of participants), 2) Permission Limiters (28.68%), 3) App Limiters (14.74%), and 4) the Privacy Unconcerned (6.84%). App and Permission Limiters were significantly more concerned about perceived surveillance than Privacy Balancers and the Privacy Unconcerned. App Limiters had the lowest number of apps installed on their devices with the lowest intention of using apps and sharing information with them, compared to Permission Limiters who had the highest number of apps installed and reported higher intention to share information with apps. The four profiles reflect the differing privacy management strategies, perceptions, and intentions of Android users that go beyond the binary decision to share or withhold information via mobile apps.

CCS CONCEPTS

• **Security and privacy → Usability in security and privacy**; • **Human-centered computing → Human computer interaction (HCI)**; • **Empirical studies in HCI**;

KEYWORDS

users profiling, smartphone users' privacy, user behaviors, privacy preferences

ACM Reference Format:
Ashwaq Alsoubai, Reza Ghaiumy Anaraky, Yao Li, Xinru Page, Bart P. Knijnenburg, and Pamela J. Wisniewski. 2022. Permission vs. App Limiters: Profiling Smartphone Users to Understand Differing Strategies for Mobile Privacy Management. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3491102.3517652

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3517652

1 INTRODUCTION

Smartphones have become an essential part of people's daily lives, holding a wealth of personal information. They are so widespread that 81% of Americans own a smartphone, and 17% of these individuals use smartphones as their primary home internet access [13]. Meanwhile, there is a high level of concern among smartphone users regarding their privacy and secondary usage of collected personal data [44, 63]. In fact, 60% of U.S. smartphone users have decided to not install certain apps due to app requests to access personal information [50]. Moreover, 46% have decided to uninstall apps upon discovering that the app collected personal information without their knowledge [50]. Alternatively, Android apps must request access to permissions for various types of information, allowing users to accept or deny requests at a granular level [28]. Controlling which privacy permissions are granted to apps decreases the risks of unwanted access to system resources or personal data [3]. Further, Google recently acknowledged how privacy permissions vary in sensitivity and classified a group of them as "Dangerous," since they give access to private user data or control over the device that can negatively impact the user [28]. Camera access is an example of a Dangerous permission as it allows apps to take photos of users and their surroundings. However, disallowing such permissions also reduces the functionality provided by apps [40]. Thus, in order to help users manage these Dangerous permissions effectively, researchers and app designers have begun to develop a better understanding of users' privacy management strategies in order to design systems that align with their privacy preferences and intentions [67].

Several privacy researchers have made strides in understanding mobile users' differing privacy attitudes, intentions, and behaviors. For instance, several privacy researchers have attempted to understand users' privacy management strategies based on their self-reported privacy attitudes or concerns (e.g., [7, 18, 39, 70]). Although self-reported privacy measures have been found to correlate to some extent with actual privacy behavior [43], researchers have also found that there is a contradiction between users' stated privacy concerns and their actual disclosure behavior, which is widely known as the "privacy paradox" [36]. Therefore, another research stream has aimed to predict and understand disclosure of private information based on users' actual privacy behavior (e.g., permissions denial behavior) [11, 41]. For example, Liu et al. profiled users based on their privacy decisions (allow or deny permissions) to

predict the app permissions settings for these profiles [41]. Yet, research has also shown that examining privacy behaviors alone also cannot fully explain or predict users' information disclosure decisions [26]. Thus, there is a need to consider both self-reported privacy attitudes and actual privacy behaviors when unpacking the complex privacy decisions of mobile phone users [26, 69]. Further, it is important to understand how mobile users' privacy attitudes and behaviors correlate to one another to ascertain if, indeed, users' privacy behaviors are paradoxical to their desired goals, or whether we as researchers may be oversimplifying the different management strategies users have devised for managing their mobile privacy. To this end, we addressed the following research questions:

RQ1: *Do users exhibit similar or differing privacy granting behaviors across the various permissions that Android has classified as "Dangerous?"*

RQ2: *How do the privacy behaviors of mobile users vary in terms of the number of apps installed on their devices versus the 'Dangerous" permissions granted to those apps?*

RQ3: *If different user profiles exist based on these privacy behaviors, how do the privacy attitudes and intentions of these groups align or differ?*

To address these research questions, we collected behavioral data scraped from user smartphones (RQ1, RQ2) and self-reported survey responses (RQ3) from Android smartphone users (N = 380). To answer RQ1, we first quantified the number of apps installed on participants' mobile device and the Dangerous permissions granted to each app. We conducted an Exploratory Factor Analysis (EFA) to reduce the 21 different types of Dangerous permissions to four stable dimensions of sensitive information disclosures pertinent for smartphone users: 1) Calendar and Contacts (Events and People), 2) Location, Camera, and Audio (Physical Information tied to the person), 3) Phone Calls (Voice communication with others), and 4) SMS and MMS (Text-based communications with others).

We then conducted a Mixture Factor Analysis (MFA) to answer RQ2 and create user profiles based on these dimensions of Dangerous Permissions and the number of apps participants had installed on their devices. We identified four unique privacy profiles based on the behavioral scraped data: 1) **Privacy Balancers** (49.74%), 2) **Permission Limiters** (28.68% of participants), 3) **App Limiters** (14.74%), and 4) **Privacy Unconcerned** (6.84%). The Privacy Balancers (largest group) demonstrated a moderate level of privacy management in terms of both app installation and permissions granting behavior. Permission Limiters (second largest group) were the least likely on average to grant permissions to apps they had installed but had the largest average number of apps installed overall. In contrast, App Limiters had the lowest average number of apps installed. Finally, the Privacy Unconcerned (smallest group) had the highest average number of Dangerous permissions granted and the second largest average number of apps installed on their devices.

To address RQ3, the differences between these profiles were examined based on the self-reported privacy attitudes and intentions (i.e., Secondary Use, Perceived Surveillance, Perceived Intrusion, Intent to Use Apps, and Intent to Share Information with Apps) of participants. We found significant differences between the profiles based on the Perceived Surveillance, Intent to Use Apps, and Intent to Share Information with Apps. For instance, App Limiters and Permission Limiters reported significantly higher levels of Perceived Surveillance than Privacy Balancers, who, in turn, reported significantly higher levels of Perceived Surveillance than the Privacy Unconcerned. App Limiters were significantly less likely to self-report a high Intention to Use Apps compared to the other three profiles. Permission Limiters had the highest Intention of Sharing Information with Apps, while App Limiters reported the lowest Intention to Share. There were no significant differences found between the profiles based on their Secondary Use and Perceived Intrusion. Our results demonstrate how smartphone users have different privacy management strategies that are fairly aligned with their self-reported privacy attitudes and intentions.

This research makes important contributions to the fields of Human-Computer Interaction (HCI) and Privacy, by presenting a new methodology to address an important and long-standing question, which involves aligning self-reported privacy preferences to ground truth behavioral data, so that we can better understand the privacy choices of mobile device users. Improving the understanding of privacy choices of mobile device users is a core contribution of this work. More specifically, this paper contributes to the networked privacy community by showing how a combination of different privacy behaviors (e.g., permissions granted and number of apps installed) can be combined to create multidimensional profiles of mobile users to characterize more complex privacy management strategies that can then be correlated with privacy attitudes and intentions. By identifying which privacy attitudes and intentions correlate with different privacy management strategies, privacy researchers and designers can better anticipate and meet the privacy goals of mobile users. Finally, by identifying four privacy profiles, we path the way for designers to account for different privacy management strategies and preferences. This can lead to better privacy management recommendations and features suited to one's privacy profile. Our work also provides practical and actionable technology and policy recommendations for the HCI community and beyond. In summary, this work makes the following novel contributions:

- Identifies unique privacy profiles for mobile users based on their multidimensional (i.e., apps installation and permission granting) privacy management strategies. We present how a mixture factor analysis can be superior to unsupervised machine learning for creating these privacy profiles through quantitative means.
- Improves the understanding of mobile users' privacy behaviors by aligning their explicit decisions to their implicit privacy attitudes and intentions, presenting evidence that given the right affordances and/or data collection and analysis methods, that the 'privacy paradox' is at best an artifact of how mobile privacy has been operationalized and measured in the past. Our work shows that mobile users find a way to behave in ways that align with with their stated cognitions.

2 BACKGROUND

We summarize two bodies of work that are relevant to our current study. First, literature having to do with mobile user privacy, which is often focused on app permissions. Second, we introduce the research that uses profiling to understand different online privacy behaviors as well as smartphone privacy behaviors.

2.1 Mobile Privacy and App Permissions

Mobile privacy has been the focus of many privacy studies exploring how users utilize the privacy features on their mobile device. Mobile app privacy is regulated by permissions granted to each application, where an app requests permission and the user either denies or grants them. Yet, Felt et al. [25] found that users struggled to comprehend the permission dialogues with 17% of their study participants paying attention to the permissions requests during installation and 3% of participants answering privacy comprehension questions correctly. Although smartphone devices offer the capability for users to protect their privacy via the permissions requests, research has shown that android users tend to grant permissions to apps "with vague descriptions and unclear purposes" [24]. Findings such as these have led researchers to investigate ways to increase users' understanding of app privacy features to better protect their privacy. For instance, one study employed nudges to help users understand the impact of granting permissions. They alert users of how often a granted permission was actually used [2]. Another approach presented by Chouhan et al. [16] allowed users to co-manage their mobile app privacy permissions within a trusted community of family and friends. Personalization is a critical component of enabling such approaches so that each mobile user can learn and be nudged about their own app activity, or set up co-managing situations appropriate for their usage [2]. Therefore, scholars have begun leveraging computational approaches to create automated solutions for permission management. These approaches predict users' permission granting behaviors (e.g., location sharing) based on their scraped behavioral data from their phones or their self-reported privacy preferences [40, 69]. An optimal solution for these computational predictions should align users' actual privacy management decisions with their privacy preferences. Relying on either scraped behavioral data or self-reported responses lacks the comprehensive understanding of how users regulate their privacy, which could affect the accuracy and real implementation of these models to predict users' privacy decisions. Therefore, this study takes a step towards understanding how well behavioral and self-reported data align with user preferences and intentions. We present an empirical approach to improve the understanding of privacy choices of mobile device users.

2.2 Profiling Users to Understand Privacy

Westin's privacy taxonomy is the foundation of this privacy profiling literature [30, 51]. In his work, he classified individuals based on their differing privacy concerns as fundamentalists, pragmatists, and unconcerned. Fundamentalists have a high privacy concern, while pragmatists have a mid-level privacy concern, and unconcerned have no or little concerns. Yet, Woodruff et al. [71] argued the validity of Westin's classification in the modern age by showing that there is no connection between Westin's privacy classification and real-world scenarios. Therefore, creating data-driven privacy profiles based on privacy behavior which are considered to be broader and more nuanced features would help to clearly understand the actual users decisions. Further, instead of viewing privacy simply as the decision to disclose or withhold information, more recent privacy researchers have moved toward profiling users based on their differing (i.e., multidimensional) privacy management strategies to better understand their behavior [39, 41, 67, 70]. For instance, Wisniewski et al. [70] were one of the first to create privacy profiles based on social media users' self-reported privacy behaviors. Their empirically grounded work identified six unique privacy profiles: Privacy Minimalists, Self-Censors, Time Savers/Consumers, Privacy Balancers, Selective Sharers, and Privacy Minimalists. They noticed that among the distinct privacy management strategies, users tend to adhere to one of these strategies, and that these strategies were associated with their privacy proficiency or awareness of privacy features. A stated limitation of Wisniewski et al.'s study was that it was based on self-reported privacy behaviors, rather than using the actual privacy settings. In the context of smartphone users, Lin et al. [39] took a similar approach to create user profiles based on self-reported privacy attitudes. They were inspired by Westin and categorized users into four groups: the unconcerned, advanced user, conservative, and fence sitters. This categorization was based on survey results about smartphone usage that asked users to rate how comfortable they were with different permissions being granted for different purposes. Our work is an extension to these previous works. Similar to Lin et al., we profile mobile users, but as an extension to this work, we designed our study so that we could model users based on actual behavioral data scraped from participants mobile devices, rather than based on their self-reports.

We also draw inspiration from Liu et al.'s work [41], where they profiled smartphone users based on their actual privacy behaviors. They analyzed privacy and security decisions of smartphone users who were asked to choose between "granting", "denying" or "requesting to be dynamically prompted" for 12 permissions of the apps they downloaded. This study demonstrated that although users' privacy behaviors are diverse, a small number of user profiles were identified with mostly similar permission management behaviors, ranging from being conservative to lenient. Liu et. al have also noted that privacy profiles can help "significantly simplifying the privacy decisions mobile user have to make" [41]. In contrast with Lin et al. [39] and Liu et al.'s [41] work, where they leveraged unsupervised machine learning, we took a different profiling approach using a Mixture Factor Analysis (MFA) when creating our profiles. The advantage of this approach is the discovery of latent dimensions resulting from the users' privacy behaviors, which represent different strategies of privacy management. MFA treats these dimensions (factors) as psychological traits (latent variables) rather than scattered behaviors. Therefore, we clustered our participants' scraped privacy decisions based on these psychological traits, which can inform us more about the psychology behind users' behaviors than pure data science [35]. Such dimensional structure can also provide a more accurate understanding of users' privacy management disclosures [35]. Thus, we make a unique contribution to the mobile privacy literature by reducing a large subset of Dangerous permissions into underlying and robust latent factors that represent different types of sensitive information disclosures made in the context of mobile app use. Then, we use these factors when profiling mobile users by their app installation and permission granting behaviors. However, users' privacy management decisions do not necessarily indicate users' intentions behind their decisions [26]. Therefore, we also go beyond prior studies by examining how our user profiles corresponded with self-reported privacy attitudes and

intentions, providing a nuanced understanding of both users' privacy decisions and intentions. We describe our methods in more detail in the next section.

3 METHODS

The goals of this study are to: 1) leverage both self-reported privacy perceptions along with actual privacy behavior to better understand the alignment between users' privacy concerns, their intentions, and their actual privacy decisions, 2) profile users based on their actual privacy behaviors, and 3) correlate the behavioral profiles to their privacy perceptions and intent. Below, we first give an overview of our study, then describe how we collected and analyzed our data.

3.1 Study Overview and Participant Recruitment

Participants were recruited through a 'Human Intelligence Task' (HIT) posted on the Amazon Mechanical Turk crowd sourcing platform. The inclusion criteria was that participants had to be adults (18 years of age or older) and live in the United States. Having high quality responses was very important to us; therefore, we limited the participants to workers who had HIT approval rates greater than 95% with having at least 50 approved HITs. The participants were required to use Android mobile devices since the application could run only on an Android operating system (v6.0 or higher).

This study was approved by our university's Institutional Review Board. MTurk users interested in taking part in the study were first directed from MTurk to a web-based explanation of informed consent to participate in the study (Appendix A). The statement of informed consent delineated important details, such as what data would be collected from their mobile device, the expected duration of the study (up to 30 minutes), and compensation for study completion. After consenting to participate in the study, participants received a consent ID and were directed to Google Play to download our study app, which required the consent ID to be entered in the app to proceed. The study took up to approximately 30 minutes for participants to complete. On average, it took participants less than 10 minutes to complete the survey portion of the study. As a confirmation of study completion, participants entered a completion code on Amazon Mechanical Turk that was provided in the app at the conclusion of the study. Participants were then asked to uninstall the app. Once participants successfully completed a study tasks, they received $1. A small number of participants who were unable to complete the study due to technical difficulties were also compensated the full amount.

In accordance with MTurk's terms of service and for our own ethical reasons, we chose to not collect any personally identifiable information from participants and took other measures to protect their personal privacy such as not collecting any use data like photos, messages, videos, voice recordings, or contacts. At the same time, we did not include the term 'privacy' anywhere in the study description to avoid priming effects. For added privacy protection, we paid participants through MTurk, which allowed us to refrain from collecting their personal email addresses. The app performed a one-time data scrape (as described in detail in the next section) and only collected necessary apps' meta-data for answering our research questions. Therefore, even if participants did not uninstall the app after study completion, no passive data collection occurred.

3.2 Data Collection

We developed an Android app that participants installed on their primary smartphone. In our analysis, we used two main measures: 1) behavioral scraped data, and 2) self-reported measures. Within the app, participants were presented with a survey where they self-reported various privacy attitudes and intentions based on pre-validated survey scales. In the background, the app scraped the Android device's applications' manifest, which is an XML file (AndroidManifest.xml) that contains meta-information about the apps installed on the device, such as app package name, package version, and permissions info including type, description, and level [29]. For this study, the data collection included the names of the apps installed on the device and the Dangerous permissions granted to each of these apps. The behavioral scraped data captured participants' actual behavior in terms of apps installation and permissions granting to apps. The self-reported measures included their Secondary Use, Perceived Surveillance, Perceived Intrusion, Intent to Use Apps, and Intent to Share Information with Apps, which are described in more detail below. This helped us gain a comprehensive understanding on how users' privacy attitudes relate to their behaviors—not only for one decision scenario, but over the course of users using different android apps. In the subsections below, we explain how we created privacy profiles based on the behavioral data (scraped from the smartphone) and examine how these profiles relate to users' self-reported data (perceived measures).

3.2.1 Scraped Behavioral Data. We scraped data from participants' app manifest to calculate the number of apps they had installed on their device and the Dangerous permissions granted to these apps.

Number of installed apps. Users' decisions to install or uninstall apps can depend on how comfortable these users to grant apps certain permissions [50]. Our decision to use apps installation behavior as a predictor for privacy behavior was based on Wisniewski et al.'s [67, 70] privacy preference and profiling work, which showed that multidimensional privacy strategies can involve indirect behaviors (e.g., unfriending, withholding information) that have subsequent privacy implications. Prior research on mobile apps found that users' concerns about their personal data being accessed by apps is one of the important factors that affects their apps' installation behaviors [34]. By installing apps, the privacy risks associated with apps requesting irrelevant permissions, vague permission definition, or permission misuse can be increased, which in turn, allow apps to access and share users' sensitive data [6, 8, 20]. These apps' malicious behaviors and malware apps affected users' decision to not install mobile apps based on their concerns about the potential unwanted access and use of personal data [19]. More recent work from Namara et al. [48] also found a significant positive correlation between the number of apps installed and the total dangerous permissions granted by a mobile user. Therefore, apps installation decisions may be considered as a behavioral indicator for users' privacy management strategies [5]. For each participant, we calculated the total number of apps installed on their device, which was used as a parameter for creating the profiles.

Granted Dangerous Permissions. We limited the permissions data to the permissions that are classified by Google as Dangerous permissions as these permissions can access the most sensitive data in a user smartphone [28]. These 21 Dangerous Permissions are listed as items in Table 4. Furthermore, we quantified permission granting behavior based on three criteria: 1) requested by the app but denied by the user, 2) requested and granted, or 3) present in the app's manifest file but not requested from the user. We calculated the permissions based on the percentage of the number of apps that had the permission "granted" over total apps that requested the permission for each participant. Next, we will describe the self-reported privacy measures we included in our analyses.

3.2.2 Mobile Users Privacy Concerns and Behavioral Intention. In this paper, we leverage the the 9-item mobile users' information privacy concerns (MUIPC) framework introduced by Xu et al. [73], which adapts Malhotra et al.'s [45] Internet Users' Information Privacy Concerns (IUIPC) scale to the mobile environment. Similar to these prior works, our measures are also based on the Theory of Planned Behavior, which states that users' intentions toward a behavior is a useful predictor for that behavior in the future [1]. This theory assumes that behavioral beliefs indicate users' intentions to perform a behavior (behavioral intention), which directly affect users' actual behavior [1]. Therefore, Xu et al. developed a multidimensional construct to quantify mobile users' privacy concerns and correlated this construct with users' behavioral intentions to use mobile apps and share private information with them. They found that privacy concern demonstrated an inverse relationship with behavioral intention, which suggested that an increase in privacy concerns led to a decrease in users' intentions to share information and use mobile apps [73]. Xu et al. also found that the dimensionality of users' information privacy concerns can be reasonably represented using three scales: Secondary Use of personal information, Perceived Surveillance, and Perceived Intrusion. They performed a survey study with (N=310) mobile users and developed the scales using Exploratory Factor Analysis (EFA) and Conformity Factor Analysis (CFA). They found that the MUIPC dimensions mediated the relationship between prior privacy experience and behavioural intention use of mobile apps.

We build upon Xu et al.'s [73] foundational mobile privacy work in two ways: first, we go beyond measuring behavioral intention towards capturing users' actual privacy behaviors in terms of installed apps and granted permissions. Second, we further explore the link between privacy concerns and behavioral intentions by developing privacy behavior profiles and linking them to users' self-reported privacy attitudes and intentions, as described in more detail in the section that follows. As in Xu et al.'s work, all constructs in our study were measured on a 5-point Likert scale (1 - Disagree Strongly. 2 - Disagree Somewhat. 3 - Neutral. 4 - Agree Somewhat. 5 - Agree Strongly.). We describe the three dimensions of mobile users' privacy concerns in more detail below.

Secondary Use of Personal Information. The Secondary Use of Personal Information construct measures individuals' concerns about the usage of the collected personal information without users' knowledge or authorization, which was originally developed by Smith et al. [59]. The Secondary Use can impact users' uncertainty about how potentially their personal information will be used in the future, placing these users in a sense of powerlessness and vulnerability [61]. This scale was adapted for the mobile privacy context and found to have a high construct validity [73]. Therefore, we included this measure to examine how mobile users actual privacy decisions might be affected by their concerns about the Secondary Use of their personal data.

Perceived Surveillance. Perceived Surveillance is the extent to which a user is concerned about having too much data collected about them without their knowledge [73]. Users' privacy decisions tend to be based on an assessment of the trade-off between the risk of being surveilled and the benefit of sharing information with apps [69]. Therefore, we included this measure as a proxy for measuring users' concern regarding the data collected from their mobile devices.

Perceived Intrusion. The Perceived Intrusion construct measures the users' intrusion perception caused by using mobile apps [73]. Solove [61] described intrusion as "invasions or incursions into one's life. It disturbs the victim's daily activities, alters her routines, destroys her solitude, and often makes her feel uncomfortable and uneasy." Therefore, we added this construct to examine its correlation with actual mobile users' decisions.

Behavioral Intention. The Theory of Planned Behavior asserts that the intention towards a given behavior can be used as a close proxy to predict that behavior [1], and that attitudes toward the behavior may in turn predict users' intentions. Secondary Use of personal information and Perceived Surveillance were both found to be significant predictors of Behavioral Intention (interestingly, Perceived Intrusion was not) [73]. Therefore, we adapted Xu et al.'s [73] MUIPC measure for Behavioral Intention in the following ways: First, we modified the time-frame from the original 12 months to 3 months. Our rationale was that app usage, and consequently the intent to disclose personal information to apps, has increased since Xu et al.'s work was published in 2012, making it more difficult to predict app use intentions over a 12-month period. Further, since participants were already agreeing to install our mobile app on their devices, it made more sense to query as to whether they would use 'new' apps during this shorter time period. Second, we added two additional items to measure participants' intention to grant location tracking permission to existing and new mobile apps. We chose location as it represents a well-studied privacy-related behavior [74] that is also considered to be a Dangerous permission for Android mobile devices.

3.3 Data Analysis Approach

For this study, we combined the analysis of self-reported privacy attitudes and intentions with scraped behavioral data from participant's mobile phones. We first validated the internal consistency of our privacy constructs, then we applied EFA for 21 Dangerous permissions to identify types of permissions that were grouped conceptually and reduce the dimensionality of the permissions. Next, the Mixture Factor Analysis (MFA) was used to create the privacy profiles based on the number of apps installed and types of the Dangerous permissions granted. The following sections describe these two data analysis approaches in more detail.

3.3.1 Evaluation of Construct Validity of Privacy Measures. We first calculated Cronbach's alpha, which measures the internal consistency of survey constructs [17]. While the Secondary Use and Perceived Intrusion satisfied commonly accepted internal consistency requirements ($\alpha = 0.91, \alpha = 0.90$ respectively), the consistency of the Perceived Surveillance ($\alpha = 0.63$) and Behavioral Intention ($\alpha = 0.65$) constructs were below the acceptable 0.7 threshold [15]. For Perceived Surveillance, the item regarding mobile users' belief of "the location of my mobile device is monitored at least part of the time" was highly rated by the majority of participants and, thus, did not correlate with the concern that mobile apps collect too much information or monitor users' activities. This speaks to the increased usage of location-based services by the majority of modern-day apps [12]. Therefore, we dropped this item, which in turn improved the internal consistency of Perceived Surveillance to $\alpha = 0.89$.

Since we added new items to measure Behavioral Intention, we performed a quick Exploratory Factor Analysis (EFA) to further examine the structure and reasons for its low internal consistency. We found that a two factor model yielded the best results, based on a principal component analysis with Varimax (orthogonal) rotation and Eigenvalues over one, explaining a total of 78.7% of the cumulative variance across all scale items. Therefore, we renamed these two scales more specifically to (as shown in Appendix B): (i) Behavioral Intention to Use Apps and (ii) Behavioral Intention to Share Information with Apps. The Behavioral Intention to Use Apps construct included two items from Xu et al.'s original work [73], while Behavioral Intention to Share Information with Apps included three items (i.e., one item from the original work of Xu et al.'s [73] and our new two location sharing items). Therefore, we divided this scale into two different types of Behavioral Intention. This analysis improved the internal consistency of the two types Behavioral Intention to ($\alpha = 0.86$) for the Intention to Use Apps and ($\alpha = 0.80$) for the Intention to Share Information with the apps. After these adjustments, we proceeded with our data analysis.

Table 1 shows the correlation between the self-reported constructs, demonstrating a significant positive correlation between the privacy concerns constructs (i.e., Secondary Use, Perceived Surveillance, Perceived Intrusion). Another correlation pattern was found between the three privacy concerns' constructs and the Behavioral Intention to Share Information with Apps which is a significant negative correlation. This suggests that mobile users who had higher privacy concerns would most likely have a lower Intention to Share Information with Apps.

3.3.2 Exploratory Factor Analysis (EFA) of Dangerous Permissions. To address RQ1, we measured users' permission granting behaviors with regard to the 21 Dangerous permissions. During our preliminary experimentation we used different data structures such as median, mean, and mode number of permissions and calculated that based on the number of apps installed per participant to discover the participants' privacy disclosure dimensionality. We conducted an Exploratory Factor Analysis (EFA) to reveal the possible underlying factors (or dimensions) of the 21 permissions. The early experimentation did not yield a good model fit based on the CFI, TLI, and RMSEA metrics. Therefore, the 21 Dangersouse permission were calculated as a percentage of the number of apps that had the permission granted over total apps that requested the permission [28] [22]. The EFA reduced the dimensionality of the 21 permission measures to four conceptual dimensions or "factors". EFA assumes a number of latent factors underlie participants' behaviors regarding the permissions, and uses the correlation values between the permissions to assign loadings to each factor [49]. The higher the value of the loading for a given permission, the more correlated the permission is with that factor. To ensure that each factor group a distinct set of highly correlated permissions, the result was then "rotated" using Geomin rotation, which is a type of of oblique rotation. Oblique rotation allows for correlations between the factors [55]. This is suitable for this study, because users' tendencies regarding different types of permissions are likely correlated (i.e., users also have an overall tendency towards disclosure). The oblique rotation allowed us to examine the correlations between the factors in a correlation matrix that is shown in the results section Table 5.

EFA is typically used to create a concise set of factors and therefore, any item (permission in our case) that did not fit these factors can be removed. Therefore, we used EFA to remove permissions that did not group well with any other permission and did not fit on any factor regardless of how many factors we would create [64] (i.e., grouping these permissions into a new factor rather than removing them would have reduced the reliability of the factor analysis [35]).

3.3.3 Mixture Factor Analysis (MFA) to Profile Users' Privacy Behaviors. To create the privacy profiles (RQ2), we used two privacy behavioral measures collected unobtrusively from the participants' devices, namely the number of applications a participant had installed, and the types of permissions the participant granted to those applications (i.e., the result of our factor analysis (RQ1)). With this data, we conducted a series of Mixture Factor Analyses (MFAs) with a robust maximum likelihood estimator. Mixture Factor Analysis is a type of factor analysis that produces clusters based on a "mixture" of factor mean scores [47]. A benefit of this approach is that it actually demonstrates how each factor (permission type) relates to permission granting behaviors for different groups of users. Studying privacy decisions with regards to the factors that behavioral permissions are based off improves the interpretability and generalizability of the findings as we study key behavioral patterns (factors) rather than many discrete behaviors [35].

Mixture Factor Analysis does not provide explicit information about the optimal number of clusters but does provide indicators that can help to compare the relative quality of solutions with differing numbers of clusters. For example, a clustering solution with a minimum value of BIC (Bayesian Information Criterion) and a maximum value for the Shannon entropy [56] (which measures "the degree of uncertainty implicated in predicting the output of a probabilistic event [57]"). can be an optimal solution [42]. In addition, while increasing the number of clusters may increase the overall fit, this increase may not be significant, so the optimal solution usually exists at the point where the log likelihood levels start to taper off [35]. These metrics may not agree on the optimal solution; therefore, one should determine the optimal number of clusters based on substantive grounds [49]: the optimal cluster solution should be based on whether the cluster distributions make sense (i.e., are interpretable) and are reasonable (e.g., avoid solutions in which certain clusters contain very few cases and/or have extreme

Table 1: The correlation between Mobile Users Privacy Concerns constructs (Secondary Use, Perceived Surveillance, Perceived Intrusion, Behavioral Intention to Use Apps, and Behavioral Intention to Share Information with Apps). This table shows that all privacy concerns constructs were significantly positively correlated together. These constructs were on the contrary significantly negatively correlated with intention to share information with the apps. * p-value < 0.05, ** < 0.01, and * < 0.001**

	SU	PS	PI	BI-UA	BI-SI
Secondary Use (SU)	1	0.741***	0.694***	-0.018	-0.198***
Perceived Surveillance (PS)		1	0.734***	0.002	-0.205***
Perceived Intrusion (PI)			1	-0.009	-0.221***
Behavioral Intention to Use Apps (BI-UA)				1	0.408***
Behavioral Intention to Share Information with Apps (BI-SI)					1

cluster means). For this study, we thus based our decision on the optimal number of clusters on this substantive grounds and the fit measures (in our case a minimum level of BIC and a maximum level of entropy).

Our MFA for privacy behavioral measures clustered participants based on their mean scores of the permission types (factors) and their number of installed applications. Table 2 compares solutions with different numbers of clusters resulting from the MFA. We did not observe any substantial improvements beyond a 4-cluster solution. The Bayesian Information Criterion (BIC), which examines the parsimony of the solution, reaches a minimum level for the 4 and 5 cluster solution. For a 4-cluster solution, the entropy reaches its maximum value, and the log likelihood levels off. Therefore, we opted to select the 4-cluster solution based on substantive grounds, backed up by the BIC, and the entropy.

Table 2: Privacy behavior MFA model fit statistics. Bold numbers indicate the best cluster solution (4 clusters) as the BIC has the minimum value and entropy has the maximum value.

Clusters	BIC	Entropy	LL
2	39117.46	0.94	-19356.8
3	39013.36	0.96	-19289.9
4	**38952.47**	**0.97**	**-19244.6**
5	38952.02	0.95	-19204.5
6	38960.47	0.95	-19168.9

Table 3: Levene's test for homogeneity of variance results. The results are significant since all p-values are less than the significance level (0.05), which indicates that there is a difference between the variances of the profiles.

Constructs	df	F	p-value
Secondary Use	3, 376	3.50	0.04
Perceived Surveillance	3, 376	7.87	<0.001
Perceived Intrusion	3, 376	0.67	0.57
Intent to Use Apps	3, 376	2.87	0.03
Intent to Share Info w/ Apps	3, 376	7.52	<0.001

3.3.4 Analysis of Variance (ANOVA) in Means of Self-Reported Privacy Perceptions. To achieve our third goal and answer RQ3, we investigated the relationship differences in the self-reported measures between the generated privacy behavioral profiles by conducting a series of Welch ANOVA tests [46] with the five self-reported measures (Secondary Use, Perceived Surveillance, Perceived Intrusion, Intent to Use Apps, and Intent to Share Info with Apps) as dependent variables and the generated profiles as the independent variable. Because Levene's test for Homogeneity of Variance was significant (see Table 3), we used the Welch ANOVA to check for significant differences between the profiles in terms of their self-reported measures. We applied a series of heteroscedastic post-hoc tests, which examines the unequal (hetero) variability across given populations, to compare individual clusters with one another [66]. The identified differences provide a comprehensive overview of the shared characteristics and distinct patterns between the behaviorally-defined privacy profiles regarding a series of privacy attitudes. The next section will display the results of our analysis including the generated permissions types, privacy profiles and uncovered differences between the privacy behavioral profiles regarding the perception and intention of the participants in these profiles.

4 RESULTS

The privacy profiles created based on the number of installed apps and the permission-granting behaviors uncover a set of distinct privacy management behaviors. The following section describes these profiles, and their distinct behavioral and attitudinal patterns.

4.1 Examining the Factors of Dangerous Permissions (RQ1)

Table 4 summarizes the underlying dimensionality (i.e., factors) of the 21 Dangerous permissions using an EFA. After removing the body sensor and processing the outgoing calls permissions due to low factor loadings (likely due to the low frequency of such permission requests), the other permissions loaded well, above a commonly used threshold of 0.4 for EFA [64]. We also highlight permissions that cross-loaded with other factors (exceeding a threshold of 0.25) in italics in Table 4. Each of the cross-loadings was substantially lower than the item's loading in its main factor. The final 4-factor model shows a good fit ($\chi^2(294) = 332.75$, p = 0.001); CFI=0.88, TLI=0.89; RMSEA=0.036, as well as good convergent validity (most

Table 4: Privacy behavioral factors and items based on EFA (4-Factors Solution). Cross-loaded permissions with other factors (exceeding a threshold of 0.25) are in italics. Removed permissions with loadings less than 0.25 are highlighted in grey. Well loaded permissions (above 0.4) for the each factor are in bold .

		Factos Loadings			
Types	**Items**	**1**	**2**	**3**	**4**
Calendar and Contacts	Read Calendar	**1.037***	0.011	-0.016	0.025
	Write Calendar	**0.751***	-0.009	0.066	-0.136*
	Read Contacts	**0.522***	0.086*	*0.289**	0.061
	Write Contacts	**0.464***	0.124*	0.125	-0.172*
Location, Camera, and Audio	Access location	-0.022	**0.821***	-0.073	0.086
	Camera	0.002	**0.959***	-0.003	0.036
	Record Audio	0.019	**0.803***	0.234*	-0.028
Removed	Body Sensors	0.048	0.034	0.099	0.030
Phone Calls	Call Phone	0.068	-0.025	**0.515***	0.006
	Read Call Log	0.010	-0.087	**0.732***	-0.027
	Read phone state	0.023	0.005	**0.618***	-0.128
	Add Voicemail	-0.037	-0.109	**0.422***	-0.203*
	Use SIP	-0.085	-0.009	**0.604***	*-0.342**
Removed	Process Outgoing Calls	-0.033	0.092	-0.038	-0.093
SMS and MMS	Send SMS	-0.004	0.057	-0.165	**0.734***
	Read SMS	0.057	-0.071	*0.464**	**0.689***
	Receive SMS	-0.048	0.024	*0.599**	**0.885***
	Receive WAP Push	0.101	0.129*	0.151*	**0.778***
	Receive MMS	0.096	0.156*	-0.016	**0.732***
	Read External Storage	-0.058	0.221*	0.470	**0.606***
	Write External Storage	0.048	0.034	0.099	**0.446***

loadings > 0.50). We labeled the four dimensions (types of permissions) of privacy behavioral measures as follows: (1) Calendar and Contacts, (2) Location, Camera, and Audio (3) Phone Calls, and (4) SMS and MMS. The first factor grouped permissions related to reading or writing from the users' Calendar and Contacts. The second permissions factor grouped permissions related to accessing users' location, camera, and recording audio. In this factor, the body sensor variable was removed because it did not have significant loadings. The body sensor permission allows access to health information such as step count, heart rate, and fitness tracker which is closely related to fitness or sports apps. We found that sports-related apps form only 0.6% of the installed apps in the participants' devices, which explains the poor performance and low loadings of the body sensor permission. The third factor was a group of permissions related to phone call, read call logs, read phone state, use SIP, and add a voicemail. In the phone calls permissions factor, the processing outgoing calls permission was removed since it did not have significant loadings. Furthermore, Table 5 shows that while all factors are significantly positively correlated with one another, the correlations are lower than the averages of the loadings per factor, indicating discriminant validity.

Table 5: Privacy permissions factor correlations. All factors are significantly positively correlated with one another. * p-value < 0.05, ** < 0.01, and * < 0.001**

1. Calendar and Contacts (CC)	1.00			
2. Location, Camera, and Audio (LCA)	0.42 *	1.00		
3. Phone calls and Voicemails (PHV)	0.35 *	0.33 *	1.00	
4. SMS & MMS (SM)	0.39 *	0.45 *	0.43 *	1.00
	CC	**LCA**	**PHV**	**SM**

4.2 Participant Profiles and Descriptive Statistics

The 380 participants in this study were nearly gender-balanced with 52% men and 48% women. Regarding their education level, the highest percentages of participants' education levels were at college (32%), a 4 years bachelor degree (25%), and 2 year college degree (Associate's) (17%). Participants had a wide variety of jobs as approximately (4%) of the participants were students, (3%) unemployed, (3%) homemakers, (3%) work on retail, and less than (1%) is distributed over a wide range of jobs including IT, analyst, and Office clerk.The scraped data shows that participants installed an average of 93.91 apps (maximum = 239, minimum = 29). The most common installed apps were: 1) tools apps (22% of the sample) such as Google Find My Device, 2) communication apps (14%) such as WhatsApp , 3) productivity apps (10%) such as Microsoft Outlook , 4) game apps (9%), and 5) entertainment apps (6%) such as Netflix.

On average, participants granted 230 Dangerous permissions to the installed apps on their devices. The most commonly granted

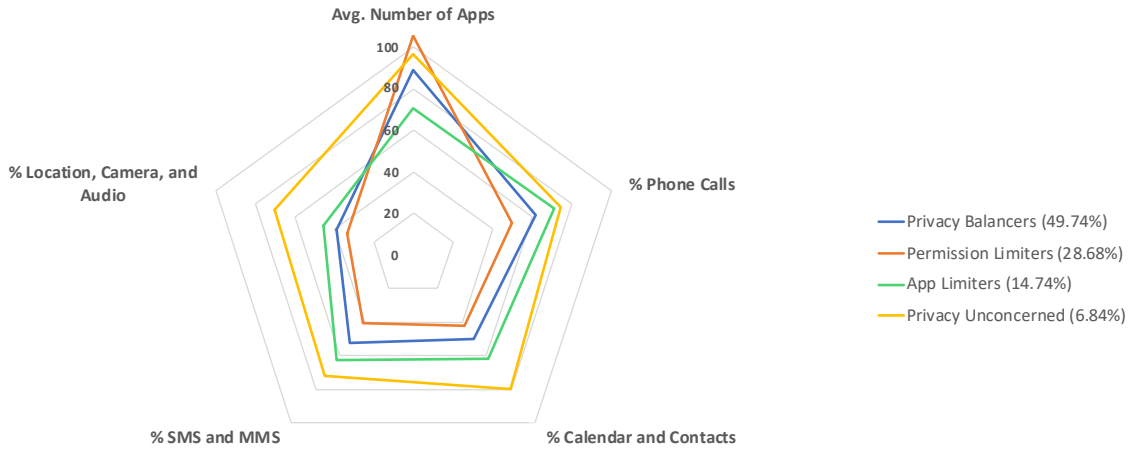

Figure 1: The four Privacy behavior profiles (N=380) based on average number of apps installed and the average number of granted permissions per type. This figure compares between the distinct privacy behavior patterns for the privacy profiles related to their apps installation and granting permissions.

Dangerous permissions included 1) Use Session Initiation Protocol (SIP), which starts a voice, video, or messaging session (27% of the total number of granted permissions), 2) Read call logs (10%), 3) Add voicemail (8%), 4) Read contacts (8%), and 5) Receive MMS (7%), while processing outgoing calls was that was the least granted Dangerous permission (0.4% of the total number of granted permissions). Apps that requested the most number of Dangerous permission were personalization (10%), maps and navigation apps (8%), and sports (8%) while music, video players, and tools requested the least number of permissions forming less than (1%) of the total number of the requested permissions. Participants accepted requested Dangerous permissions on average 20% of the time but were most likely to accept for tools (48%), communications (43%), and travel apps (34%). While games (4%), personalization (4%), education (3%), and sports (2%) had the lowest accept ratios.

4.3 Privacy Behavior Profiles (RQ2)

To answer RQ2, we classified users into distinct profiles based on their actual apps installed and types of permission granted. The resulting four MFA clusters represent a set of distinct "privacy profiles" that describe the members of each cluster. The average number of installed apps and the average number of granted permissions per type for each profile are shown in Figure 1. This figure shows the multidimensionality of privacy management strategies based on apps installation and permissions granting decisions for the generated privacy profiles. We labeled the four privacy profiles as follows:

4.3.1 Privacy Balancers (49.74%): The Privacy Balancers practiced a moderate level of privacy management for both behaviors (apps installation and permission granting). In comparison with other profiles, Privacy Balancers installed a moderate number of apps (m = 88) and granted a moderate level of permissions for all permissions factors. Overall, they installed more apps on average than the App Limiters but less apps than the Permission Limiters and the Privacy Unconcerned. In all four categories of Dangerous permissions, they granted more permissions on average than Permission Limiters but fewer than the App Limiter and Privacy Unconcerned. This profile was representative of the largest user group, comprised of nearly half of all participants.

4.3.2 Permission Limiters (28.68% of participants): Permission Limiters had the highest number of installed apps on their devices on average (m = 105). However, they granted the least number of permissions for all categories of Dangerous permissions in comparison to other profiles. Permission Limiters managed their privacy by limiting access to all sensitive information on their devices and granting the fewest Dangerous permissions to the large number of installed apps they had. Since Permission Limiters managed their privacy through permission management, this strategy may make them feel more confident about installing more apps than the other three profiles.

4.3.3 App Limiters (14.74%): Among the generated privacy profiles, App Limiters installed the least number of apps on average (m = 70). Yet, we found that App Limiters have moderately high levels of granting permission behavior, granting more permissions than Privacy Balancers and Permission Limiters. This difference suggests that App Limiters have a different approach to safeguarding their privacy; they take a selective app installation approach and limit the total number of applications on their devices. Once they install apps, they are less concerned about restricting permissions granted to those apps.

4.3.4 Privacy Unconcerned (6.84%): The Privacy Unconcerned showed generous granting behavior for all permissions and had the second highest average number of apps installed on their devices (m = 98). This included frequently granting permissions for apps to access their camera, location, and audio at significantly higher

Table 6: Mean and standard deviation of the Secondary Use, Perceived Surveillance, Perceived Intrusion, Intent to Use Apps , and Intent to Share Info with Apps measures by the privacy profiles.

	Secondary Use		Perceived Surveillance		Perceived Intrusion		Intent to Use Apps		Intent to Share Info w/ Apps	
Privacy Profiles	M	SD	M	SD	M	SD	M	SD	M	SD
Privacy Balancers	3.64	0.38	4.29	0.68	4.01	0.38	3.78	1.00	4.38	0.82
Permission Limiters	3.70	0.34	4.50	0.43	4.12	0.34	4.25	0.60	4.59	0.62
App Limiters	3.68	0.49	4.59	0.41	3.98	0.38	2.14	0.63	3.75	1.00
Privacy Unconcerned	3.61	0.32	3.55	0.48	4.09	0.32	4.06	0.58	4.36	0.85

rates than all other privacy profiles. These Dangerous permissions are arguably the most sensitive in terms of giving visual, audio, and physical access to an individual and their environment, as opposed to digital access to the device itself. Compared to the other three profiles, this group appears to be unconcerned about their privacy protection, since they generously grant Dangerous permissions requested to a larger percentage of apps installed on their devices. The next section will examine these profiles in relation to participants' self-reported privacy attitudes and intentions.

4.4 Examining Differences in Privacy Attitudes and Intentions between Profiles (RQ3)

In this section, we explain how participants' self-reported privacy attitudes align with their actual disclosure behaviors. We analyzed participants' self-reported privacy attitudes and intentions . Means and standard deviations are listed in Table 6. The Welch ANOVA results presented in Table 7 show a significant difference between the privacy profiles in terms of Perceived Surveillance, Intent to Use Apps, and Intent to Share Info with Apps. In the following subsections, we discuss these differences in the self-reported measures between the profiles.

4.4.1 *Mobile Users Privacy Concerns.*

Secondary Use. An ANOVA did not yield any significant difference between the profiles regarding their self-reported Secondary Use concern ($F(3, 93.31) = 0.39, p = 0.75$) as shown in Table 7. Table 6 showed the privacy profiles' mean scores to the Secondary Use construct (Privacy Balancers: $m = 3.64$, Permission Limiters: $m = 3.70$, App Limiters: $m = 3.65$, and App Limiter: $m = 3.61$). This suggested that all profiles fell in the range of being "Neutral" to "Agree Somewhat" (based on the Likert anchors) to the Secondary Use measure.

Perceived Surveillance. An ANOVA revealed significant differences between the four profiles ($F(3, 98.64) = 34.17, p < 0.001$) based on the Perceived Surveillance measure. Post-hoc tests (Table 8) demonstrated that App Limiters ($m = 4.59$) and Permission Limiters ($m = 4.50$) perceived a significantly higher level of surveillance than Privacy Balancers ($m = 4.29$), and that the Privacy Unconcerned ($m = 3.55$) perceived a significantly lower level of surveillance than all other groups (see Table 9). Generally, these results align well with the four profiles. Since the Privacy Unconcerned are less worried about Perceived Surveillance, for instance, they are willing to install more apps and grant more Dangerous permissions to those apps. In contrast, the App and Permission Limiters are significantly more concerned about surveillance, so take specific measures (e.g., limiting apps or permissions) to protect their personal privacy.

Perceived Intrusion. Regarding Perceived Intrusion, there were no significant differences found between the profiles based on the ANOVA test ($F(3, 94.18) = 0.55, p = 0.64$) as listed in table Table 7. Looking at the average scores of these profiles in Table 6, it showed that privacy profiles shared similar high level of privacy intrusion concern (Privacy Balancers: $m = 4.01$, Permission Limiters: $m = 4.12$, App Limiters: $m = 3.98$, and App Limiter: $m = 4.09$). These mean scores showed that all privacy profiles mostly selected "Somewhat agree" on the scale for the Perceived Intrusion construct.

Table 7: Welch ANOVA results. The significant results are in bold (p-values less than 0.05), suggesting that there are significant differences between the privacy profiles based on these constructs. Table 8 and Table 9 unpack the found significant differences based on the Post-hoc tests.

Constructs	df	F	p-value
Secondary Use	3,93.31	0.39	0.75
Perceived Surveillance	**3, 98.64**	**34.17**	**<0.001**
Perceived Intrusion	3, 94.18	0.55	0.64
Intent to Use Apps	**3, 89.79**	**11.17**	**<0.001**
Intent to Share Info w/ Apps	**3, 101.07**	**142.72**	**<0.001**

4.4.2 *Intent to Use Apps.*

We also studied the self-reported scores of participants' Intent to Use Apps. Means and standard deviations are listed in Table 6. We found significant differences between the four profiles ($F(3, 89.79) = 11.17, p < 0.001$). Post-hoc tests (Table 8) revealed that in accordance with their behavior, App Limiters ($m = 2.14$) expressed a significantly lower intention to use apps than Privacy Balancers ($m = 3.78$), Permission Limiters ($m = 4.25$), and the Privacy Unconcerned ($m = 4.06$) (see Table 9). Overall, we found that participants' intention to use apps aligned well with their behavioral profiles. Particularly, App Limiters, who had the lowest number of apps installed, expressed the lowest intention to use apps.

4.4.3 *Intent to Share Information with Apps.*

We also studied participants' self-reported Intent to Share Info with Apps. Means and standard deviations are again listed in Table 6. We found significant differences between profiles in regard to Intent to Share Info with Apps ($F(3, 101.07) = 142.72, p < 0.001$). Post-hoc tests (Table 8)

Table 8: The significant differences between the profiles based on the Post-hoc test.

Constructs	Privacy Profiles	95% Confidence Interval
Perceived Surveillance	App Limiters > Privacy Balancers	[0.09791, 0.49292]
	App Limiters > Privacy Unconcerned	[0.74316, 1.34476]
	Permission Limiters > Privacy Balancers	[0.03400, 0.37552]
	Permission Limiters > Privacy Unconcerned	[0.66776, 1.23885]
	Privacy Balancers > Privacy Unconcerned	[0.45440, 1.04268]
Intent to Use Apps	App Limiters < Permission Limiters	[-1.23533, -0.44816]
	App Limiters < Privacy Balancers	[-1.02487, -0.23703]
	App Limiters < Privacy Unconcerned	[-1.19969, -0.03108]
Intent to Share Info with Apps	App Limiters < Permission Limiters	[-2.38305, -1.83309]
	App Limiters < Privacy Balancers	[-1.93648, -1.34262]
	App Limiters < Privacy Unconcerned	[-2.30750, -1.52309]
	Permission Limiters > Privacy Balancers	[0.22339, 0.71365]

Table 9: Summary of significant pairwise differences.

Privacy Constructs	Significant Pairwise Differences (Mean)
Perceived Surveillance	App Limiters (4.59), Permission Limiters (4.50) > Privacy Balancers (4.29) > Privacy Unconcerned (3.55)
Intent to Use Apps	Permission Limiters (4.25), Privacy Unconcerned (4.06), Privacy Balancers (3.78) > App Limiters (2.14)
Intent to Share Info with Apps	Permission Limiters (4.59), Privacy Balancers (4.38), Privacy Unconcerned (4.36) > App Limiters (3.75)
	Permission Limiters (4.59) > Privacy Balancers (4.38)

showed that App Limiters (m = 3.75) expressed a significantly lower intention to share information with apps than than Privacy Balancers (m = 4.38), Permission Limiters (m = 4.59), and the Privacy Unconcerned (m = 4.36). We also found a significant difference as Permission Limiters (m = 4.59) had a higher intentions to share info with apps more than the Privacy Balancers (m = 4.38). (see Table 9). In this case, it also makes sense that App Limiters, who limit the number of apps installed on their devices, would also intend to limit the amount of information they shared with apps. Yet, it seems somewhat paradoxical that Permission Limiters would intend to share higher levels of information with apps than the Privacy Balancers. A possible explanation for this finding is that since Permission Limiters take granular measures to limit permissions they do not want certain apps to have access to, they are more inclined to share other kinds of information (e.g., less sensitive) information with these apps.

5 DISCUSSION

In the sections below, we first discuss the importance of the factor structure of Dangerous permissions. Then, we discuss the implications of the four privacy profiles and how this relates to self-reported privacy perceptions.

5.1 Grouping Dangerous Permissions into Meaningful Factors

Although Google classifies all of the Dangerous permissions into a single high-risk group [23], our results emphasize that users do not treat all Dangerous permissions the same way. Similar to existing work [35, 67], participants in our study treat different types of information as contextually and semantically different. Applying EFA to the Dangerous permissions allowed us to uncover this multidimensionality for Android Dangerous permissions. Namely, we identified four latent dimensions of Dangerous permissions: 1) Calendar and Contacts, 2) Location, Camera, and Audio, 3) Phone calls, and 4) SMS and MMS. As peoples' privacy management behavior can be broken down into distinct factors, one could argue for a permissions interface that would allow users to grant permissions on a group-by-group basis rather than considering each individual permission separately. Our EFA found that permissions within a given group (e.g., Phone Calls) are often treated in the same way by users. Leveraging this knowledge could simplify the process of granting semantically similar groups of permissions. For instance, default privacy settings can be set per group of rather than per permission, e.g. Android could simply deny the Location, Camera, and Audio permissions by default. Likewise, grouping these categories conceptually within privacy permission dialogues may be more user-friendly than asking for permissions one by one. It might also

reduce cognitive load by not making privacy decisions too complex for less tech savvy end users [25, 32].

At the same time, considering permissions on a per-group basis could also potentially pose an additional data exposure threat [10]. For instance, the "Calendar and Contacts" group of permissions could be considered related to People and Events—an email app could be given these permissions, enabling the app to easily retrieve the contact email addresses and give users the ability to share or send scheduled events to potential attendees. However, if a malicious apps would take advantage of this bundled set of permissions, it would receive sensitive information that is not only related to the mobile users' schedule but also to the people in their contact list. Similarly, the group of Location, Camera, and Audio can be very helpful in case of emergencies (e.g., an SOS app that sends this info to an emergency contact in the case of emergency), it could also allow apps to misuse the permissions (e.g, stalking purposes). Based on our results, the Location, Camera, and Audio permissions were the least likely to be granted by all profiles (see Figure 1)—arguably, participants found them riskier because they bridge the boundary between virtual and physical surveillance [14, 33].

The multidimensionality we found in the android permissions makes a methodological contribution regarding how to measure privacy permission behaviors. The majority of prior work has measured users' privacy permission behavior by using an itemized list of various permission items and treating each item separately [24, 40], rather than treating permissions as semantically meaningful constructs. A downside to this approach is that it does not consider the relationships among the permission items and ignores these underlying regularities in users' permission granting behaviors. Future research can use our Dangerous permissions categorization to gain a more holistic understanding of users' permission control decisions and and to quantify the effects of granting the different groups of Dangerous permissions on the overall mobile user experience and privacy.

These findings also have important design implications for privacy researchers and app designers as they need to be aware of how certain specific Dangerous permissions may be perceived as particularly sensitive. For example, when designing new privacy features which require a Dangerous permission, app developers may consider where the permission fits across the four identified permission types (see Table 4) and consider how different privacy profiles may view the sensitivity of this type of disclosure. For instance, most of our participants (except the Privacy Unconcerned) were least likely to grant permissions when they involved a sensor that gathered external data (e.g., location, camera, and audio) beyond the phone. Meanwhile, they were more likely to share call-related information (e.g., phone calls) and phone-based information (e.g., Calendar and Contact). As such, app developers should do their due diligence before requesting the most sensitive of these Dangerous permissions. Similarly, when considering whether to ask users to grant permissions, designers should be sure that there is a clear and high value justification before trying to gain access to smartphone users' Dangerous permission categories [72]. In summary, both researchers and practitioners can gain valuable insights from our work. Our approach could be used as an exemplar for others who wish to conduct similar types of studies or build privacy-centered and user-centric mobile interfaces.

5.2 Profiling Mobile Users by App Installation and Permission Granting Behavior

Clustering participants in our study along two types of privacy behaviors (i.e., number of apps installed and granting of Dangerous permissions) allowed us to identify four distinct privacy profiles: 1) Privacy Balancers, 2) Permission Limiters, 3) App Limiters, and 4) Privacy Unconcerned. Our study is different from previous studies that create privacy profiles derived from self-reported preferences [30, 40]; instead, our classification was derived directly from the users' actual privacy management behaviors (i.e., scraped data). A benefit of our approach is that we could scrape smartphone users' privacy settings, categorize them into one of the four profiles, and make immediate recommendations on how they can best manage their privacy (e.g. by automatically supporting users' privacy decisions related to the apps and permissions configurations), thereby avoiding the "cold start" problem (i.e., the need for extensive external input from users) of most computer-based recommendation systems [53]. Further, our profiling approach resulted in profiles that have empirically validated relationships to users' self-reported privacy attitudes and intentions to ensure that any intelligent defaults or recommendations made would align with users' stated privacy goals as well as their privacy behavioral profiles.

Two of the profiles uncovered in our work demonstrated distinctly opposite privacy management strategies: participants in one profile restricted the number of apps but generously granting permissions, while participants in the other profile installed a large number of apps but restricting the permissions granted to these apps. Our study suggests that these differences in behaviors trace back to differences in privacy concerns and/or intentions, which we discuss further in the next section. These different behaviors have design implications for how to assist users in their privacy management. For instance, it may be helpful to use a background service to automatically detect apps that have not been used for a while. Once an unused app is identified, the system might recommend uninstalling the app to an App Limiter, while it might suggest removing the app's access to Dangerous permissions to a Permission Limiter. A third option would be to utilize the cloud to "offload Unused Apps" (see iOS), thereby *temporarily* removing apps that a user did not use for a while [58]. This could allow users with the habit of installing a large amount of apps (e.g., Permissions Limiters), to more effectively and thoroughly decide on the right amount of apps on their devices and avoid unnecessary data leakage. Regarding the generous permission granting behavior in the case of the Unconcerned and Apps Limiters profiles, an analysis of how often a given app access a certain resource could be utilized to provide recommendations as to whether a certain permission may be revoked without sacrificing the functionality of the apps. This type of recommendation could better assist users in managing their permissions more effectively, rather than indiscriminately granting all or even denying all permissions.

5.3 The Importance of Examining Behavioral Attitudes and Actual Privacy Behaviors

While there is a large body of literature on understanding and improving permission dialogues to align the choices of users with

their privacy preferences [21, 36, 48, 69], our approach of combining app installations and permission settings with self-reported privacy attitudes and intentions of mobile users is a novel contribution of our work. By considering app installation behaviors in conjunction with permission granting behaviors and mapping these behaviors into privacy profiles, we were able to connect these privacy decisions back to users' privacy attitudes and intentions, which gave us a more holistic understanding of how users' privacy goals were actualized in their observed behavior. Our results demonstrate that users' privacy attitudes and behaviors are somewhat consistent—While previous research on the "privacy paradox" has shown conflicting results between users' privacy concerns and their actual behavior [36], we were able to find an alignment of the privacy management profiles with users' stated privacy attitudes and intentions. Thus, a key finding of this study is that the actual behavior expressed by the privacy profiles aligned fairly well with the self-reported privacy attitudes. For instance, participants who were attitudinally more privacy-cautious (i.e., those who scored higher on the Perceived Surveillance scale, lower on the Intent to Share Info with Apps scale, and lower on the Intent to Use Apps scale), also tended to utilize privacy preserving strategies, such as having fewer apps installed or fewer Dangerous permissions granted to their apps.

Of the three dimensions of mobile users' privacy concern, Perceived Surveillance was significantly different between the profiles, while Secondary Use and Perceived Intrusion were non-significant. When Xu et al. presented the MUIPC framework, they found Perceived Surveillance had the strongest effect size in terms of predicting Behavioral Intention [73]. Our findings confirm that Perceived Surveillance was also the most impactful in terms of differentiating between our privacy profiles based on actual behavior. Therefore, our paper adds to the mobile privacy literature by showing that Perceived Surveillance is still the most useful concern in predicting mobile users' actual privacy decisions. This demonstrates an important modern-day mobile users' concern about the continuous surveillance practices by mobile apps that aggressively collect their data such as identities, daily behaviors, or their real-time locations [38]. Our results for RQ3 also suggest that, while Secondary Use and Perceived Intrusion are significantly and positively correlated with Behavioral Intention to Share Information with apps (Table 1), they may not be as useful in predicting actual mobile privacy behaviors. This may be partly due to the changing norms and expectations about Secondary Use (which on average ranged from "Neutral" to "Agree Somewhat" across all of our participants) and Perceived Intrusion, of which most participants "Agreed Somewhat" with this mobile privacy concern, that have converged among Android smartphone users over time. This raises a question about the feasibility of continuing to measure such constructs in assessing modern-day mobile users' privacy concerns and behavioral intentions when trying to understand actual privacy behavior. As a lesson learned, future privacy research should carry out efforts to continually evaluate the psychometric value of pre-validated privacy constructs to reflect the realities of the time.

Further, our results show that it is possible that end users may employ different strategies to regulate their privacy beyond traditional privacy settings. For example, some social network users manage their privacy by creating multiple accounts within the same social network site for different circles of friends, even though this is not explicitly considered a privacy behavior [68]. Similarly, a decision such as not installing or uninstalling an app could be done for privacy reasons but may not be considered an explicit privacy behavior. A recent study indeed demonstrates an impact of privacy risk perceptions on users' decision to install apps [62]. These results, in combination with the ones discussed in this paper, provide compelling evidence that future privacy research should go beyond predefined privacy settings or disclosure behaviors to adopt a broader perspective on different privacy strategies that could contribute to a more holistic understanding of how mobile users actually tailor their privacy.

Having a nuanced understanding of privacy profiles, which are based on the users' behaviors, is critical in the process of building more personalized privacy experiences. Recently, there has been a focus on leveraging machine learning algorithms to build recommendation systems that predict users' privacy preferences [4, 31, 54, 65, 69]. These systems were developed and trained using either past privacy decisions [65] or self-reported privacy preferences [60]. Further research should incorporate both of these aspects, as users' privacy decisions alone cannot be considered as an adequate predictor of their privacy preferences. The incorporation of both past behavior and current preferences into a recommender system can help build more accurate privacy preference recommendation systems. A practical example could be a smartphone that gives the users the ability to choose a privacy management profile based on a trained privacy preference recommendation system (cf. [54]). This profile could help the system to tailor the user experience to the user's privacy expectations. The Privacy Unconcerned, for instance, would likely be more open to recommendations for new apps they might want to install on their device, while Permission Limiters might value a tool that continually scans their apps for permissions that could be revoked or limited based on use. Likewise, Privacy Balancers may benefit from "privacy cost" versus "app benefit" dialogues that allow them to engage in a process of weighing various trade-offs through a privacy calculus [27, 37] for installing new apps and/or granting Dangerous permissions.

6 LIMITATIONS AND FUTURE WORK

Our findings are specific to our sample, which included Android users over the age of 18 who were recruited on Amazon Mechanical Turk. Past research has shown that participants recruited via Amazon Mechanical Turk demographically approximate the U.S. population [9] and are generalizable to the general populations' security and privacy experiences [52]. We also checked the latest US Census for 18 years and over to compare the demographics between our sample and US Census. Regarding gender, both MTurk and US Census have almost a balance distribution between women and men. We also found similar patterns of races between the two samples. However, there were some slight differences regarding education and income. For education level, our sample had a higher number of people who attended some college or had obtained an associate's or bachelor's degree than the U.S. Census. Our sample also had a higher percentage of low income participants compared to the U.S. Census. As such, our results may not be generalizable to mobile users with different demographic backgrounds than those

who participated in our study or iOS smartphones users. Future research should extend this line of inquiry to more representative populations or to different groups, such as younger or older smartphone users or those in different education levels. Future studies could also conduct deeper investigations of permission granting behaviors in relation to different types of apps. Also, given the lower income of MTurk workers found in our sample, we strongly recommend that future studies (including our own) that recruit MTurk workers pay them a living wage for their time and effort.

Another limitation of our study was its cross-sectional design, where we did not measure usage patterns or behavior over time. For instance, we were unable to capture how often our participants installed and uninstalled apps over time. Given the results of our study, we believe that this is an important behavior that should be measured as it has important implications for one's mobile privacy. In all cases, researchers should take into account that multiple privacy behaviors might work in concert towards a user's privacy management goals. Understanding different approaches by measuring multiple privacy behaviors as well as attitudes can help distinguish between seemingly paradoxical attitudes when taken in aggregate.

7 CONCLUSION

Understanding how users manage their privacy allows researchers and designers to better anticipate their needs as they design smartphone apps and anticipate user response. In this paper, we presented four privacy profiles based on user behavior which shed light on how individuals approach smartphone app privacy management in different ways. Users balance app installation with granting Dangerous permissions in different ways to achieve their desired level of privacy. By anticipating a variety of privacy management behaviors and styles, researchers can better understand users and map their intentions to behaviors.

ACKNOWLEDGMENTS

This work was partially supported by a grant from the Bentley Data Innovation Network and partially supported by the U.S. National Science Foundation (NSF) grant number #CNS-1814439.

Conflict of interest statement. Any opinion, findings, recommendations, and conclusions expressed in this material are solely those of the authors and do not necessarily reflect the views of the Bentley Data Innovation Network or the U.S. National Science Foundation.

REFERENCES

[1] Icek Ajzen. 1991. The theory of planned behavior. *Organizational behavior and human decision processes* 50, 2 (1991), 179–211.

[2] Hazim Almuhimedi, Florian Schaub, Norman Sadeh, Idris Adjerid, Alessandro Acquisti, Joshua Gluck, Lorrie Faith Cranor, and Yuvraj Agarwal. 2015. Your location has been shared 5,398 times! A field study on mobile app privacy nudging. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, Seoul, Korea, 787–796.

[3] Rawan Baalous and Ronald Poet. 2018. How Dangerous Permissions are Described in Android Apps' Privacy Policies?. In *Proceedings of the 11th International Conference on Security of Information and Networks*. ACM, New York, United States, 1–2.

[4] Paritosh Bahirat, Yangyang He, Abhilash Menon, and Bart Knijnenburg. 2018. A Data-Driven Approach to Developing IoT Privacy-Setting Interfaces. In *23rd International Conference on Intelligent User Interfaces (IUI '18)*. ACM, New York, NY, USA, 165–176. https://doi.org/10.1145/3172944.3172982

[5] Susanne Barth and Menno DT De Jong. 2017. The privacy paradox–Investigating discrepancies between expressed privacy concerns and actual online behavior–A systematic literature review. *Telematics and informatics* 34, 7 (2017), 1038–1058.

[6] Susanne Barth, Menno DT de Jong, Marianne Junger, Pieter H Hartel, and Janina C Roppelt. 2019. Putting the privacy paradox to the test: Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources. *Telematics and informatics* 41 (2019), 55–69.

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

[8] Christoph Buck, Chris Horbel, Tim Kessler, and Claas Christian. 2014. Mobile consumer apps: Big data brother is watching you. *Marketing Review St. Gallen* 31, 1 (2014), 26–35.

[9] Martin J Burnham, Yen K Le, and Ralph L Piedmont. 2018. Who is Mturk? Personal characteristics and sample consistency of these online workers. *Mental Health, Religion & Culture* 21, 9-10 (2018), 934–944.

[10] Paolo Calciati, Konstantin Kuznetsov, Alessandra Gorla, and Andreas Zeller. 2020. Automatically Granted Permissions in Android apps: An Empirical Study on their Prevalence and on the Potential Threats for Privacy. In *Proceedings of the 17th International Conference on Mining Software Repositories*. ACM, Seoul, Korea, 114–124.

[11] Weicheng Cao, Chunqiu Xia, Sai Teja Peddinti, David Lie, Nina Taft, and Lisa M. Austin. 2021. A Large Scale Study of User Behavior, Expectations and Engagement with Android Permissions. In *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Virtual, 803–820. https://www.usenix.org/conference/usenixsecurity21/presentation/cao-weicheng

[12] Pew Research Center. 2016. Americans increasingly use smartphones for more than voice calls, texting. https://www.pewresearch.org/internet/ft_01-27-16_smartphoneactivities_640/

[13] Pew Research Center. 2021. Demographics of Mobile Device Ownership and Adoption in the United States. https://www.pewresearch.org/internet/fact-sheet/mobile/. Retrieved May 9, 2021.

[14] Saksham Chitkara, Nishad Gothoskar, Suhas Harish, Jason I Hong, and Yuvraj Agarwal. 2017. 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 (2017), 1–22.

[15] Eunseong Cho and Seonghoon Kim. 2015. Cronbach's coefficient alpha: Well known but poorly understood. *Organizational research methods* 18, 2 (2015), 207–230.

[16] Chhaya Chouhan, Christy M LaPerriere, Zaina Aljallad, Jess Kropczynski, Heather Lipford, and Pamela J Wisniewski. 2019. Co-designing for community oversight: Helping people make privacy and security decisions together. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–31.

[17] Lee J Cronbach and Paul E Meehl. 1955. Construct validity in psychological tests. *Psychological bulletin* 52, 4 (1955), 281.

[18] Prajit Kumar Das, Anupam Joshi, and Tim Finin. 2017. Personalizing context-aware access control on mobile platforms. In *2017 IEEE 3rd International Conference on Collaboration and Internet Computing (CIC)*. IEEE, IEEE, San Jose, CA, USA, 107–116.

[19] Larry Dignan. 2011. Google's Android wears big bullseye for mobile malware. https://www.zdnet.com/article/googles-android-wears-big-bullseye-for-mobile-malware/

[20] Manuel Egele, Christopher Kruegel, Engin Kirda, and Giovanni Vigna. 2011. PiOS: Detecting Privacy Leaks in iOS Applications. In *Proceedings of the Network and Distributed System Security Symposium, NDSS 2011, 6th February - 9th February 2011*. The Internet Society, San Diego, California, USA, 15. https://www.ndss-symposium.org/ndss2011/pios-detecting-privacy-leaks-ios-applications-paper

[21] Serge Egelman, Adrienne Porter Felt, and David Wagner. 2013. *Choice architecture and smartphone privacy: There'sa price for that*. Springer, Germany. 211–236 pages.

[22] Leandre R Fabrigar and Duane T Wegener. 2011. *Exploratory factor analysis*. Oxford University Press, England.

[23] Zheran Fang, Weili Han, and Yingjiu Li. 2014. Permission based Android security: Issues and countermeasures. *computers & security* 43 (2014), 205–218.

[24] Johannes Feichtner and Stefan Gruber. 2020. Understanding privacy awareness in android app descriptions using deep learning. In *Proceedings of the Tenth ACM Conference on Data and Application Security and Privacy*. ACM, New Orleans, LA, USA, 203–214.

[25] Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. 2012. Android permissions: User attention, comprehension, and behavior. In *Proceedings of the eighth symposium on usable privacy and security*. ACM, Washington, D.C., 1–14.

[26] Nina Gerber, Paul Gerber, and Melanie Volkamer. 2018. Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & security* 77 (2018), 226–261.

[27] Reza Ghaiumy Anaraky, Kaileigh Angela Byrne, Pamela J Wisniewski, Xinru Page, and Bart Knijnenburg. 2021. To Disclose or Not to Disclose: Examining the

Privacy Decision-Making Processes of Older vs. Younger Adults. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Virtual, 1–14.

[28] Google. 2021. Google and Open Handset Alliance. n.d. Android API Guide. https://developer.android.com/guide/topics/permissions/overview#dangerous_permissions. Retrieved Feb 21, 2021.

[29] Google. 2022. Android Developers. https://developer.android.com/guide/topics/manifest/manifest-element

[30] L Harris, AF Westin, et al. 2003. Consumer Privacy Attitudes: A Major Shift Since 2000 and Why.

[31] Yangyang He, Paritosh Bahirat, Bart P. Knijnenburg, and Abhilash Menon. 2019. A Data-Driven Approach to Designing for Privacy in Household IoT. *ACM Trans. Interact. Intell. Syst.* 10, 1 (Sept. 2019), 10:1–10:47. https://doi.org/10.1145/3241378

[32] Patrick Gage Kelley, Sunny Consolvo, Lorrie Faith Cranor, Jaeyeon Jung, Norman Sadeh, and David Wetherall. 2012. A conundrum of permissions: installing applications on an android smartphone. In *International conference on financial cryptography and data security*. Springer, Kralendijk, Caribbean Netherlands, 68–79.

[33] Asma Khatoon and Peter Corcoran. 2017. Android permission system and user privacy—a review of concept and approaches. In *2017 IEEE 7th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)*. IEEE, Berlin, 153–158.

[34] Jennifer King, Airi Lampinen, and Alex Smolen. 2011. Privacy: Is there an app for that?. In *Proceedings of the Seventh Symposium on Usable Privacy and Security*. ACM, Pittsburgh, Pennsylvania, USA, 1–20.

[35] Bart P Knijnenburg, Alfred Kobsa, and Hongxia Jin. 2013. Dimensionality of information disclosure behavior. *International Journal of Human-Computer Studies* 71, 12 (2013), 1144–1162.

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

[37] Robert S Laufer and Maxine Wolfe. 1977. Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of social Issues* 33, 3 (1977), 22–42.

[38] Christian Fernando Libaque-Sáenz, Siew Fan Wong, Younghoon Chang, and Edgardo R Bravo. 2021. The effect of fair information practices and data collection methods on privacy-related behaviors: a study of Mobile apps. *Information & Management* 58, 1 (2021), 103284.

[39] Jialiu Lin, Bin Liu, Norman Sadeh, and Jason I Hong. 2014. Modeling users' mobile app privacy preferences: Restoring usability in a sea of permission settings. In *10th Symposium On Usable Privacy and Security ({SOUPS} 2014)*. Usenix, Menlo Park, California, 199–212.

[40] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun Aerin Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. 2016. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016)*. usenix, Denver, Colorado, USA, 27–41.

[41] Bin Liu, Jialiu Lin, and Norman Sadeh. 2014. Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help?. In *Proceedings of the 23rd international conference on World wide web*. ACM, Seoul, Republic of Korea, 201–212.

[42] Gitta H Lubke and Bengt Muthén. 2005. Investigating population heterogeneity with factor mixture models. *Psychological methods* 10, 1 (2005), 21.

[43] Christoph Lutz and Pepe Strathoff. 2014. Privacy concerns and online behavior–Not so paradoxical after all? Viewing the privacy paradox through different theoretical lenses. *Viewing the Privacy Paradox Through Different Theoretical Lenses (April 15, 2014)* 4 (2014), 81–99.

[44] Mary Madden, Lee Rainie, Kathryn Zickuhr, Maeve Duggan, and Aaron Smith. 2014. Public perceptions of privacy and security in the post-Snowden era.

[45] Naresh K Malhotra, Sung S Kim, and James Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research* 15, 4 (2004), 336–355.

[46] Karl Moder. 2010. Alternatives to F-test in one way ANOVA in case of heterogeneity of variances (a simulation study). *Psychological Test and Assessment Modeling* 52, 4 (2010), 343–353.

[47] Bengt Muthén and Bengt O Muthén. 2009. *Statistical analysis with latent variables*. Wiley, New York, NY.

[48] Moses Namara, Reza Ghaiumy Anaraky, Pamela Wisniewski, Xinru Page, and Bart P Knijnenburg. 2021. Examining Power Use and the Privacy Paradox between Intention vs. Actual Use of Mobile Applications. In *European Symposium on Usable Security 2021*. ACM, Virtual, 223–235.

[49] Karen L Nylund, Tihomir Asparouhov, and Bengt O Muthén. 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural equation modeling: A multidisciplinary Journal* 14, 4 (2007), 535–569.

[50] K Olmstead and M Atkinson. 2017. Apps permissions in the Google Play store. Pew Research Center.

[51] Privacy and American Business. 1997. Commerce, Communication and Privacy Online: A National Survey of Computer Users.

[52] Elissa M Redmiles, Sean Kross, and Michelle L Mazurek. 2019. How well do my results generalize? comparing security and privacy survey results from mturk, web, and telephone samples. In *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, San Francisco, CA, USA, 1326–1343.

[53] Neil Rubens, Mehdi Elahi, Masashi Sugiyama, and Dain Kaplan. 2015. Active learning in recommender systems. In *Recommender systems handbook*. Springer, Boston, MA, 809–846.

[54] Odnan Ref Sanchez, Ilaria Torre, Yangyang He, and Bart P. Knijnenburg. 2019. A recommendation approach for user privacy preferences in the fitness domain. *User Modeling and User-Adapted Interaction* 30 (Oct. 2019), 513–565. https://doi.org/10.1007/s11257-019-09246-3

[55] Thomas A Schmitt. 2011. Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment* 29, 4 (2011), 304–321.

[56] Claude Elwood Shannon. 2001. A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review* 5, 1 (2001), 3–55.

[57] Guey-Shin Shyu, Bai-You Cheng, Chi-Ting Chiang, Pei-Hsuan Yao, and Tsun-Kuo Chang. 2011. Applying factor analysis combined with kriging and information entropy theory for mapping and evaluating the stability of groundwater quality variation in Taiwan. *International Journal of Environmental Research and Public Health* 8, 4 (2011), 1084–1109.

[58] Drew Smith. 2020. iOS Client Administration. In *Apple macOS and iOS System Administration*. Springer, Apress, 109–144.

[59] H Jeff Smith, Sandra J Milberg, and Sandra J Burke. 1996. Information privacy: Measuring individuals' concerns about organizational practices. *MIS quarterly* 20 (1996), 167–196.

[60] Daniel Smullen, Yuanyuan Feng, Shikun Zhang, and Norman M Sadeh. 2020. The Best of Both Worlds: Mitigating Trade-offs Between Accuracy and User Burden in Capturing Mobile App Privacy Preferences. *Proc. Priv. Enhancing Technol.* 2020, 1 (2020), 195–215.

[61] Daniel J Solove. 2005. A taxonomy of privacy. *U. Pa. L. Rev.* 154 (2005), 477.

[62] Siok Wah Tay, Pin Shen Teh, and Stephen J Payne. 2021. Reasoning about privacy in mobile application install decisions: Risk perception and framing. *International Journal of Human-Computer Studies* 145 (2021), 102517.

[63] TRUSTe. 2014. *US consumer confidence privacy report: consumer opinion and business impact*. Technical Report. Research Report, TRUSTe Inc.

[64] George Ursachi, Ioana Alexandra Horodnic, and Adriana Zait. 2015. How reliable are measurement scales? External factors with indirect influence on reliability estimators. *Procedia Economics and Finance* 20 (2015), 679–686.

[65] Primal Wijesekera, Arjun Baokar, Lynn Tsai, Joel Reardon, Serge Egelman, David Wagner, and Konstantin Beznosov. 2017. The feasibility of dynamically granted permissions: Aligning mobile privacy with user preferences. In *2017 IEEE Symposium on Security and Privacy (SP)*. IEEE, San Jose, CA, USA, 1077–1093.

[66] Rand R Wilcox. 2011. *Introduction to robust estimation and hypothesis testing*. Academic press, Global.

[67] Pamela Wisniewski, AKM Islam, Heather Richter Lipford, and David C Wilson. 2016. Framing and measuring multi-dimensional interpersonal privacy preferences of social networking site users. *Communications of the Association for information systems* 38, 1 (2016), 10.

[68] Pamela Wisniewski, Heather Lipford, and David Wilson. 2012. Fighting for my space: Coping mechanisms for SNS boundary regulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin, Texas, USA, 609–618.

[69] Pamela Wisniewski, Muhammad Irtaza Safi, Sameer Patil, and Xinru Page. 2020. Predicting smartphone location-sharing decisions through self-reflection on past privacy behavior. *Journal of Cybersecurity* 6, 1 (2020), tyaa014.

[70] Pamela J Wisniewski, Bart P Knijnenburg, and Heather Richter Lipford. 2017. Making privacy personal: Profiling social network users to inform privacy education and nudging. *International Journal of human-computer studies* 98 (2017), 95–108.

[71] Allison Woodruff, Vasyl Pihur, Sunny Consolvo, Laura Brandimarte, and Alessandro Acquisti. 2014. Would a privacy fundamentalist sell their DNA for $1000 ... if nothing bad happened as a result? the westin categories, behavioral intentions, and consequences. In *10th Symposium On Usable Privacy and Security ({SOUPS} 2014)*. Usenix, Menlo Park, California, 1–18.

[72] Zhiqiang Wu, Xin Chen, Muhammad Umair Khan, and Scott Uk-Jin Lee. 2021. Enhancing Fidelity of Description in Android Apps With Category-Based Common Permissions. *IEEE Access* 9 (2021), 105493–105505.

[73] Heng Xu, Sumeet Gupta, Mary Beth Rosson, and John M Carroll. 2012. Measuring mobile users' concerns for information privacy. In *Thirty Third International Conference on Information Systems*. Citeseer, Orlando, FL, USA, 16.

[74] Heng Xu and Hock-Hai Teo. 2004. Alleviating consumers' privacy concerns in location-based services: a psychological control perspective. In *Proceedings of the International Conference on Information Systems*. Association for Information Systems, Washington, DC, USA, 64.

A STUDY INFORMED CONSENT

Thank you for agreeing to participate in our research. Before you begin, please note that the data you provide may be collected and used by Amazon as per its privacy agreement. This agreement shall be interpreted according to United States law.

A.1 Why am I being invited to take part in a research study?

We invite you to take part in a research study because you stated that you are an Android smartphone user in the United States. You must be 18 years of age or older to participate in this study.

A.2 What should I know about a research study?

- Someone will explain this research study to you.
- Whether or not you take part is up to you.
- You can choose not to take part.
- You can agree to take part and later change your mind.
- Your decision will not be held against you.
- You can ask all the questions you want before you decide.

A.3 Why is this research being done?

The purpose of this research is to understand what permissions users grant or deny for applications installed on their Android smartphones. This a better understanding of this behavior, we may be able to improve the design of permission default settings to enhance the user experience for future Android smartphones.

A.4 How long will the research last?

The study should take no longer than 30 minutes to complete.

A.5 What happens if I say yes, I want to be in this research?

After consenting to participate in this study, you will be asked to download and install an Android app from the Google Play store. Upon opening the app, it will collect and analyze app permissions from your phone. For example, whether or not the Facebook app was given permission to track your location. You will also be asked to answer survey questions via the app. All data collected is anonymous, meaning that you cannot be personally identified by it in any way (even by us). We will NOT have access to any personal data from your phone, including photos, messages, videos, voice recordings, and contacts.

A.6 What happens if I do not want to be in this research?

Participation in research is completely voluntary. You can decide to participate or not to participate.

A.7 What happens if I say yes, but I change my mind later?

You can leave the research at any time, and it will not be held against you. If you decide to leave the research, you will not get paid for the MTurk HIT.

A.8 What happens to the information collected for the research?

Efforts will be made to limit the use and disclosure of your personal information, to people who have a need to review this information. No personally identifiable information will be collected that would allow the researchers or others to identify who you are. However, we cannot promise complete secrecy. Organizations that may inspect and copy your information include the IRB and other representatives of this organization.

A.9 What else do I need to know?

Once you complete the survey and your data passes standard quality checks, you will receive a code which you can use to get paid one U.S. dollar ($1.00) via Amazon Mechanical Turk. Please delete the research app after you are finished.

To proceed with this study, please agree to the terms by selecting the appropriate response below.

"I Agree", "I Disagree"

B SURVEY QUESTIONS

B.1 Self-reported Privacy Measures

The following self-reported constructs were adapted from Xu et al. [73] and measured on a 5-point Likert Scale: Strongly Disagree - Strongly Agree.

Construct	Questions
Secondary Use of Personal Information	I am concerned that mobile apps may use my personal information for other purposes without notifying me or getting my authorization.
	When I give personal information to use mobile apps, I am concerned that apps may use my information for other purposes.
	I am concerned that mobile apps may share my personal information with other entities without getting my authorization.
Perceived Surveillance	I believe that the location of my mobile device is monitored at least part of the time.
	I am concerned that mobile apps are collecting too much information about me.
	I am concerned that mobile apps may monitor my activities on my mobile device.
Perceived Intrusion	I feel that as a result of my using mobile apps, others know about me more than I am comfortable with.
	I believe that as a result of my using mobile apps, information about me that I consider private is now more readily available to others than I would want.
	I feel that as a result of using mobile apps, information about me is out there that, if used, will invade my privacy.
Behavioral Intention to Use Apps	I predict I will use new mobile apps in the next 3 months.
	I intend to use mobile apps in the next 3 months.
Behavioral Intention to Share Information with Apps	I am likely to disclose my personal information to use mobile apps in the next 3 months.
	I am likely to grant permission to share my location with my existing mobile apps in the next 3 months.
	I am likely to grant permission to share my location with new mobile apps in the next 3 months.

B.2 Demographic Information

Questions	Possible Responses
How many years have you lived in the United States?	• I do not live in the U.S. • Less than 1 Year • 1 to 3 Years • 4 to 5 Years • Longer Than 5 Years
Which state in the U.S do you live in?	(select state)
How would you characterize the locality you live in?	• Urban • Suburban • Rural
What is the highest level of education you have completed?	• Less than high school diploma • High school diploma • Some college • 2 year college degree (Associate's) • 4 year college degree (Bachelor's) • Some graduate school • Master's degree • Doctoral degree (PhD) • Professional Degree (MD, JD, MBA)
What is your sex? [1]	• Male • Female • Other • Do not wish to specify
What is your ethnic background? (Select all that apply.)	• White/Caucasian • Black/African-American • Hispanic/Latino • Asian • Native Hawaiian or Other Pacific Islander • American Indian/Alaska Native • Other, (please specify)
What is your current employment status? (Select all that apply.)	• Employed full time • Employed Part time • Unemployed looking for work • Unemployed not looking for work • Homemaker • Student • Retired • Disabled • Other , (please specify)
What is your occupation?	(open-response)

[1] https://interactions.acm.org/archive/view/july-august-2019/how-to-do-better-with-gender-on-surveys