JOURNAL OF CYBERSECURITY

Journal of Cybersecurity, 2020, 1–16
doi: 10.1093/cybsec/tyaa014
Research paper

Research paper

Predicting smartphone location-sharing decisions through self-reflection on past privacy behavior

Pamela Wisniewski,[1,*] Muhammad Irtaza Safi,[1,†] Sameer Patil[2] and Xinru Page[3]

[1]College of Engineering and Computer Science, University of Central Florida, Orlando, FL 32816, USA; [2]Luddy School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN 47408, USA and [3]Department of Computer Science, Brigham Young University, Waltham, Provo, UT 84602, USA

*Correspondence address: 4000 Central Florida Blvd., Harris Corporation Engineering Center (HEC) 217A, Orlando, FL 32816, USA. E-mail: pamwis@ucf.edu

[†]Master's Student at College of Engineering and Computer Science, University of Central Florida, Orlando, FL, 32816, USA

Received 31 May 2019; revised 9 February 2020; accepted 19 May 2020

Abstract

Smartphone location sharing is a particularly sensitive type of information disclosure that has implications for users' digital privacy and security as well as their physical safety. To understand and predict location disclosure behavior, we developed an Android app that scraped metadata from users' phones, asked them to grant the location-sharing permission to the app, and administered a survey. We compared the effectiveness of using self-report measures commonly used in the social sciences, behavioral data collected from users' mobile phones, and a new type of measure that we developed, representing a hybrid of self-report and behavioral data to contextualize users' attitudes toward their past location-sharing behaviors. This new type of measure is based on a reflective learning paradigm where individuals reflect on past behavior to inform future behavior. Based on data from 380 Android smartphone users, we found that the best predictors of whether participants granted the location-sharing permission to our app were: behavioral intention to share information with apps, the "FYI" communication style, and one of our new hybrid measures asking users whether they were comfortable sharing location with apps currently installed on their smartphones. Our novel, hybrid construct of self-reflection on past behavior significantly improves predictive power and shows the importance of combining social science and computational science approaches for improving the prediction of users' privacy behaviors. Further, when assessing the construct validity of the Behavioral Intention construct drawn from previous location-sharing research, our data showed a clear distinction between two different types of Behavioral Intention: self-reported intention to use mobile apps versus the intention to share information with these apps. This finding suggests that users desire the ability to use mobile apps without being required to share sensitive information, such as their location. These results have important implications for cybersecurity research and system design to meet users' location-sharing privacy needs.

Introduction

Mobile users now account for the majority of Internet traffic (52%) [1], and mobile app revenue is projected to reach $188 billion by 2020 [2]. However, the wealth of personal information that mobile apps access has a considerable impact on their usage. According to a recent Pew study on smartphone app usage in the USA, when deciding to install an app, 90% of smartphone owners find it important to know how their personal information will be used [3]. In fact, 60% of app users reported deciding not to install an app because of the personal information it requested, and 46% had uninstalled an app after discovering the extent to which it collected personal information [3]. To encourage app adoption and use, researchers and app designers need to be able to understand and predict people's willingness to disclose various types of personal information toapps and use this knowledge to help align the system with users' privacy needs and expectations [4].

However, predicting online personal information disclosure is not straightforward. Research has uncovered discrepancies between people's stated concerns and their actual disclosure behaviors. This widely acknowledged "privacy paradox" [5, 6] has made it difficult to predict user behavior based on stated privacy preferences. Instead, social science researchers have developed scales to predict and explain users' privacy intentions [7–10]. Yet, for the most part, there is still a gap when mapping these self-reported measures to actual behavior. Meanwhile, computational scientists typically make predictions based on behavioral data, such as the number of apps installed on a device, the kinds of apps installed, and the types of permissions granted to the apps. However, computational studies have typically used this behavioral data to detect problematic behavior or to alert users regarding potential data leaks [11–13]. The collected data has generally been used to predict behavior as opposed to helping users reflect on their past privacy behaviors and inform or guide their future privacy decisions. The goal of the present research is to examine whether self-reflection on past privacy behavior can help improve the prediction of a person's future privacy decisions. In turn, this improved predictive power and understanding can help researchers and designers better meet the privacy preferences of end-users [14].

In this study, we focused on understanding and predicting location-sharing behavior in the context of mobile apps. We compared the effectiveness of self-reported measures, behavioral data, and newly developed measures of self-reflection on past privacy behavior to answer the following research questions:

RQ1: *What factors best predict users' app location-sharing behavior for three types of predictor variables?*

A. pre-validated self-reported constructs from the literature,
B. scraped behavioral data from smartphones, and
C. hybrid measures capturing users' perceptions of their past location-sharing behavior based on the data scraped from their smartphones.

RQ2: Across the three types of predictor variables above, what combination of factors best predicts users' app location-sharing behavior?

To this end, we conducted a study with 380 Android users recruited via Amazon Mechanical Turk. We developed an Android app that participants installed on their smartphones. The app measured self-reported constructs via an embedded questionnaire embedded and unobtrusively scraped behavioral data, such as the number of apps with permission to access location. Then, the app asked participants whether they were comfortable with the current location permissions for each of the apps installed on their phones (a construct we refer to as a self-reported measure in the context of past behavior). We asked participants to share their location with our app and treated their sharing decisions as our dependent variable.

For RQ1A, we leveraged several pre-validated self-reported constructs from the literature that have been shown to be relevant for predicting behavioral intent toward location-sharing: Behavior Intention [15], Perceived Surveillance [10], Perceived Intrusion [16], Secondary Use of Personal Information [17], "For Your Information" (FYI) communication style [8], and Power Usage [7]. Participant responses to these pre-validated constructs were collected via a questionnaire embedded in the study app. To investigate RQ1B, we scraped behavioral data in the background, including the manifest of the apps installed on the phone, the 'Dangerous Permissions' [18] granted to these apps, and specifically, whether location permissions were granted to these apps. According to Google, Dangerous Permissions are those that provide access to a user's personal information or stored data or control an app's access to the operation of other apps [18]. To answer RQ1C, we created our own measures by showing participants the existing apps installed on their smartphones that had been granted location-sharing permissions. We asked participants to reflect on their comfort levels sharing their location with each of these apps and to indicate whether they would revoke this permission. To analyze this data, we carried out binary logistic regression in a step-by-step fashion. This allowed us to derive the strongest model for predicting whether participants granted the location permission to our app (RQ2).

Overall, we found that self-reported measures and our new hybrid measure of self-reflection on past privacy behaviors were the strongest predictors of location-sharing behavior. The model that used self-reported variables (RQ1A) explained 16.5% of the variance in location-sharing behavior. Behavioral Intention to Share Information with Apps and FYI Communication Style were the significant predictors in this model. The scraped behavioral data (RQ1B) explained only 3.2% of the variance in the model with Number of Installed Apps as the only significant predictor. The model that included only hybrid measures of self-reflecting on past privacy behaviors explained 12.4% of the variance in our dependent variable. Of the three hybrid measures, Location Comfort (the percentage of apps installed on the participants' smartphones that they were comfortable having access to their location) was the only significant predictor of location-sharing behavior. The combined model with both generic self-reported measures and self-reflection measures in the context of past behavior explained 19.5% of the variance in whether participants shared their location with our app. The difference in explanatory power between these models was statistically significant, confirming that the addition of our new hybrid variable improved the overall predictive value of the final model.

Improving the ability to predict people's location-sharing behaviors can help us understand the factors that underlie users' privacy decisions. In addition, app developers can anticipate user attitudes and create a better user experience by asking for the location permission only when users are likely to grant it. Overall, we make the following contributions to the field of end-user mobile privacy and, in general, cybersecurity; we: (i) identify the generic self-reported measures which best predict users' actual location-sharing behavior, (ii) show that scraped behavioral data is not a good predictor of location-sharing behavior, and (iii) uncover that a hybrid self-

reported measure of users reflecting on their past behavior can significantly improve models for predicting location-sharing behaviors.

Related work

Studies of privacy on mobile devices have generally been divided into "social" and "computational" approaches. Social science researchers typically conduct survey studies with self-reported measures to gauge user behavior whereas computational researchers often rely on behavioral data scraped from the device to predict user privacy decisions. In the sections that follow, we describe these two approaches and argue for the need of merging the two in order to advance interdisciplinary research on understanding and predicting privacy behavior.

Leveraging attitudes to predict behavior

Research on location-sharing through technologies often focuses on the connection between privacy concerns and attitudes and location disclosure behavior. Therefore, we draw on the body of literature that measures privacy attitudes to predict location disclosure. In addition, we look at other relevant attitudes that affect location-sharing behavior.

Measuring privacy attitudes: Survey-based studies on mobile privacy have been conducted mainly in the fields of Information Systems (IS) and Human-Computer Interaction (HCI). A key theme of these studies has been to try to predict user behavior based on self-reported survey responses. Such questionnaires generally try to measure the user's attitude toward specific topics using specially designed constructs (for a review of various information privacy measures, see Preibusch [19]). For instance, a commonly accepted practice in such research is that user behavior can be predicted based on user attitudes and that Behavioral Intention (based on the Theory of Planned Behavior [20]) is the strongest predictor of actual behavior. The Theory of Planned Behavior states that behavioral beliefs inform user attitudes toward behaviors which then lead to behavioral intention that directly impacts user behavior [20]. Therefore, researchers have developed various constructs to quantify different aspects of user beliefs, feelings, intentions, and attitudes toward mobile privacy in order to predict behavioral intention as a dependent variable.

Smith *et al.*'s [17] work was one of the first to develop scales to measure user concern for information privacy. The work introduced a fifteen-item instrument that measured concerns regarding data collection, unauthorized secondary use, improper access, and errors in data handling. Malhotra *et al.* [21] extended this scale and adapted it for online privacy contexts to create the Internet Users' Information Privacy Concerns scale (IUIPC). Similarly, Buchannan [22] developed an online privacy scale that differentiates between General Caution and Technical Protection. Our research focuses on the context of mobile privacy. Therefore, we drew from the Mobile Users' Information Privacy Concern (MUIPC) framework [10] that identifies factors that influence mobile phone users' behavioral intention to use mobile apps and share their personal information with the apps. MUIPC is a three-factor scale that determines users' concern for information privacy in a mobile context by measuring their concern regarding Misuse of Shared Data, Degree of Intrusion, and Perceived Surveillance. Since MUIPC has been employed in several user behavior studies (e.g., [23, 24]), we drew heavily from prevalidated items that were originally derived from this scale and adapted it for mobile location sharing.

Other attitudes related to location-sharing: In this research, we further drew from several studies that specifically examined users' location-sharing behaviors. For example, Guha *et al.* [25] studied the practice of deceptive location sharing (i.e., users deliberately sharing incorrect location information) and found that users engaged in this practice as acts of boundary and impression management due to various concerns about privacy. Page *et al.* [8] showed that location-sharing decisions were influenced by a specific communication style called "For Your Information" (FYI), where users would rather infer availability and social information about others than interact and ask forthis information explicitly. Those preferring an FYI Communication style were more likely to use location-sharing social networks. Further work by Page *et al.* [26] found that the desire to preserve relationship boundaries was the main source of privacy concerns regarding location-sharing social networks; when people felt that location sharing would change their relationships with others, they expressed a wide range of social privacy concerns, such as worrying about feeling compelled to interact with others or being inundated with information from other people. Heavy users of location-sharing social media were found to be less concerned that location sharing could impact their relationship boundaries. Other location-sharing studies reinforce the potential privacy concerns induced by sharing location. Barkhuus *et al.* [27] distinguished between location-tracking services (i.e., services that disclose a user's location to others) and position-aware services (i.e., services that rely on the device's knowledge of its location). Even though users perceived both of these services as equally useful, Barkhuus *et al.* [27] found that location-tracking services produced far greater concern for privacy.

Additionally, researchers have described power users as those who use technology to the fullest extent, adapt easily to technological changes, and feel that technology is an integral part of their lives [7]. Power users were found to prefer customization (i.e., tailoring interfaces to match their preferences) as long as the customization was self-initiated. If the customization was done automatically by the system, power users felt a loss of agency and did not feel as positively toward the presented content [9]. Therefore, we incorporated Power Usage as a construct in our model, which is described in more detail in our Research Framework.

Computational approaches to mobile privacy

In contrast, computational approaches to mobile privacy generally analyze the data generated as a result of the system operation and user interaction with the software. These analyses typically apply machine-learning techniques for a number of purposes, such as segmenting the population based on privacy behaviors [28], identifying malware [12], and detecting and surfacing potentially privacy invasive operations [11]. In addition, computational approaches for privacy support have analyzed other data sources, such as the application source code [13], operating system stack traces [29], and network traffic [30]. In turn, many of these research endeavors attempt to raise user awareness regarding potential privacy violations and empower and influence users to make privacy-preserving choices. For instance, Li *et al.*'s [13] PERUIM tool presented users with a graphical representation of the mapping between an application's user interface and permissions. Almuhimedi *et al.*'s [11] permissions manager sent users periodic 'nudges' depicting data collection practices of apps installed on their phones, resulting in 58% of the participants restricting some of their permissions. Fu *et al.* [31] similarly found that runtime location access disclosures surfaced unexpected accesses and led users to make privacy-preserving adjustments to their settings.

Further, computational analyses have been applied to reduce the burden of privacy management by determining the appropriate preferences on the user's behalf. For instance, Liu *et al.* [32] implemented a Personalized Privacy Assistant that makes personalized recommendations for privacy preferences, and Fawaz *et al.* [33] built LP Guardian to provide location privacy with minimal user interaction. Research has found that incorporating contextual awareness can improve the quality of predicted user preferences [34, 35]. The number and kinds of permissions requested by an app have also been shown to signal privacy and security risks. As a result, analyses of app permissions have been employed to detect malicious apps [36, 37] and assess app privacy risks [38]. Research has shown that many commonly used dangerous permissions could be replaced by less granular permissions that preserve privacy with little to no loss of functionality and overhead [39].

While these computational approaches facilitate detecting privacy and security risks, raising user awareness, and helping users consider privacy risks, they are less useful in helping researchers and designers understand the reasons behind privacy preferences and behavior. More importantly, in these approaches, predictions of user behavior rely on past behavior being a reliable indicator of the future, which may not necessarily be the case. Moreover, computational efforts typically assume that user behavior is well-aligned with privacy preferences. However, the privacy paradox calls this assumption into question. As a result, computational approaches may potentially reinforce past mistakes or nudge users toward privacy decisions that they might regret [40].

An interdisciplinary approach to mobile privacy

There is a relative scarcity of research that has merged the above two approaches by collecting survey data and simultaneously scraping behavioral data to try to understand and predict user privacy behavior. In one study, Lin et al. [41] downloaded apps through the Google play store and performed static analysis of the app source code to identify specific sections within the code where permissions were used. Separately, they recruited participants to answer questions about the downloaded apps, using these responses to derive a set of privacy profiles based on correlating self-reported preferences with app permissions within the source code. Lin *et al.* [42] found that clearly revealing the purpose of requesting sensitive permissions made it more likely that users felt positively about granting permissions to an app. Similarly, Ghosh *et al.* [43] used phone metadata, such as call frequency and call length, to predict user privacy concerns. They found that higher call response rate, higher missed call rate, and higher number of new contacts were associated with a low concern for privacy.

These studies demonstrate the potential benefit of adopting interdisciplinary approaches to understand and predict privacy behaviors and/or attitudes. Information collected from users can further serve as ground truth to enable verification of various predictions based on the data scraped from devices. For instance, researchers have shown that the list of apps installed on a person's phone can predict the person's demographic characteristics and other personal attributes with high levels of accuracy [44–46]. While Lin *et al.* [41] created privacy profiles for users and Ghosh *et al.* [43] predicted privacy concerns, we used self-reported measures, scraped behavioral data, and hybrid measures capturing users' perceptions about their past behavior to predict user privacy behavior (specifically, whether participants chose to grant location access to our study app). In contrast with Lin *et al.* [42], who asked participants general questions regarding apps, our hybrid measures allowed us to tailor our questions to the specific apps installed on the participants' devices. For example, instead of asking participants whether they are comfortable sharing their location with mobile apps in general, we scraped the app manifest of all apps installed on their mobile phone and asked specifically about location-sharing comfort regarding each app to which they had already granted location access. We consider such contextually prompted responses as a type of self-reflection on past behavior. We believe this new type of variable represents a synergistic construct that combines the strengths of the social and computational sciences in a way that helps us better understand end-user privacy behavior.

Research framework

Our goal was to predict whether users would share their location with our app based on the following distinct classes of independent variables: (i) generic self-reported measures, (ii) scraped behavioral data, and (iii) hybrid measures capturing users' perceptions of their past location-sharing behavior in the context of reflecting on the data scraped from their smartphones. To ensure a baseline understanding of the key differences among these measures, we provide a brief introduction to each.

Self-reported constructs measure how a user feels about a certain topic, behavior, or action. Self-reported measures involve self-reported user data that is grouped into sets of related items pertaining to the underlying constructs. Constructs are often considered latent variables, i.e., variables that are inferred from other observed variables. A construct must be carefully designed to be statistically valid [47], i.e., it must pass various construct validity tests so that it is confirmed to be measuring what it claims to measure. For the purposes of our study, we used pre-validated constructs from prior research specific to mobile location sharing and/or privacy research. An example of a pre-validated construct that we leveraged for our study is the Behavioral Intention construct used by Xu *et al.* [15] to quantify the degree to which users plan on disclosing their location information and using mobile apps. They asked users their intent "to use mobile apps," and disclose "personal information to use mobile apps" in the next 12 months. These statements were asked using a 5-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree."

In contrast, behavioral measures are content or metadata that is collected unobtrusively from users' devices. This data is considered "objective" in that it is based on scraped data as opposed to the users' subjective self-reports. Studies by computer scientists often use objective variables since such data can be collected via software. An example of using scraped behavioral data to make predictions is Seneviratne *et al.*'s [44] work predicting user traits, such as religion, gender, and relationship status, based on the kinds of apps installed on the phone.

We propose a third category of measures, inspired by the education literature showing that a reflective learning approach produces positive learning outcomes [48]. This approach consists of an individual reflecting on past choices to inform future decisions [49]. In light of previous missteps and poor choices, individuals are more likely to make better choices in the future. We created hybrid measures that follow this approach by combining behavioral data to add context to self-reported measures. First, participants were shown their behavioral data scraped from their smartphones and then asked to reflect on those past choices. The self-reported measures consisted of asking their self-reflection on that past behavior and their future intentions on whether to rectify the choices. We tested

this novel class of measures to examine their effectiveness in predicting location-sharing behavior.

In the subsections below, we describe our dependent variable and explain how we operationalized the three classes of predictor variables, including our new hybrid measures.

Dependent variable: location sharing

Location services are the cornerstone of personalized mobile content and play an important role in delivering targeted information and advertisements. Social networks, such as Foursquare, and mapping applications, such as Google Maps, rely on users sharing their locations with the community or the app. Therefore, finding models to improve the prediction of whether users would share their location can go a long way in helping such applications design better user experiences. However, the literature points to the sensitivity of sharing such data. Much location-sharing research has been devoted to understanding with whom and under what conditions people are willing to share their location [50–53]. A number of studies have attempted to predict end-user privacy behavior in mobile contexts. For instance, research has linked attitudes and behavior regarding granting permissions with the clarity of describing the purpose of the requested access [42, 54]. Others have focused on the recipients of the location information and their relationships with the sharer [26, 52, 55]. Still other research has found that location sharing could undermine trust in family relationships [56].

If software designers can anticipate privacy concerns associated with sharing location, they can design better user experiences that meet users' privacy expectations and improve user retention. Therefore, we chose our dependent variable as a measure of privacy behavior representing whether a person agrees to provide location access to our study app. We framed the variable as a 'Yes/No' choice based on whether the location permission was granted.

Self-reported measures

For self-reported measures, we included relevant constructs from the mobile and location disclosure literature. The constructs are described in more detail in Appendix 1. All measures were used in their original form, except for the Behavior Intention construct, which we describe in depth in the next section.

Behavior Intention: We adapted this measure from Xu *et al.* [15] who drew from the earlier work of Malhotra *et al.* [21] and their Internet Users' Information Privacy Concerns (IUIPC) scale. The construct Behavioral Intention hails from the Theory of Planned Behavior which states that behavior can be predicted using attitudes toward the behavior and behavioral intent [20]. Xu *et al.* used Behavioral Intention as the dependent variable in their research to show that increased privacy concern reduced a user's behavioral intention to disclose personal information and use mobile apps [15]. We slightly modified the wording of the items by shortening the intention-to-use time frame from 12 months down to 3 months. We felt that a time frame of 3 months would be more realistic for the context of our study because the adoption and use of mobile apps has increased significantly since Xu *et al.*'s paper was published in 2012, and participants in our study would have just installed our mobile app to participate in our study. We created two additional behavioral intention items specific to location sharing: (i) "I am likely to grant permission to share my location with my existing mobile apps in the next 3 months," and (ii) "I am likely to grant permission to share my location with new mobile apps in the next 3 months." We added these new items to the existing ones in the scale.

Perceived Surveillance: Perceived Surveillance is a construct developed and validated by Xu *et al.* [10] to quantify users' perception of being surveilled and having too much information collected about them. Perceived Surveillance is one of the factors in MUIPC [10] and is rooted in the dimension of "collection." Malhotra *et al.* [21] noted that data collection is the starting point of various privacy concerns. Therefore, the dimension of collection in IUIPC measures the degree to which a person is concerned about disclosing specific data relative to the value of benefits received. Since user privacy decisions are often based on an assessment of perceived benefits and risks associated with the decision, we added the Perceived Surveillance measure to quantify the perception of the balance between surveillance and benefit [57].

Perceived Intrusion: Developed in Xu *et al.*'s earlier work [16], Perceived Intrusion quantifies the perception of intrusion caused by using mobile apps. Xu et al. [16] conducted a survey and found that perceived intrusion shaped people's views about the privacy practices of specific websites. Therefore, we used the perceived intrusion measure in order to correlate it with the privacy-related decision of sharing location. Our intent was to examine if the perceived intrusion of mobile app use influenced location-sharing behavior.

Secondary Use of Personal Information: The Secondary Use of Personal Information construct quantifies people's concerns about their information being used for purposes other than those for which it was collected [17]. Solove *et al.* [58] noted that "the potential for secondary use generates fear and uncertainty over how one's information will be used in the future, creating a sense of powerlessness and vulnerability." Similarlly, Xu *et al.* [10] included secondary use as a factor in their MUIPC scale. Therefore, we included this measure in order to study the impact of the perception of secondary use on location-sharing behavior.

FYI Communication Style (About Myself): The FYI Communication Style was identified as an online communication-style preference by Page *et al.* [8]. It consists of two subscales, one quantifying the extent to which one prefers to communicate one's own location (FYI About Myself) and the other representing preferences for learning about others' location in an FYI way (FYI About Others). People with the FYI Communication Style tend to keep in touch with others without direct interaction and prefer to have location shared broadly. These individuals were shown to be more willing to share their location in location-sharing social networks [8]. FYI About Myself and FYI About Others were shown to be highly correlated, and only one is needed to predict behavior. Since the FYI About Myself construct has already been shown to impact location-sharing decisions, we included it to represent attitudes toward sharing one's location.

Power Usage: The Power Usage scale from Marathe *et al.* [7] measures the degree to which someone is a "power user." Power users are technologically adept and use their devices to the fullest potential. Kang *et al.* [59] found that power users are less likely to share personal information on personalized mobile sites but reveal more when interacting with non-personalized mobile content. Since our study app is highly personalized, we used this measure to analyze whether being a power user affects location-sharing choices.

Scraped behavioral variables

We collected the following data from the devices of our study participants:

Number of Installed Apps: The Number of apps installed on the device is the apps installed on the users' device as included in the device app manifest. We added this variable because it directly corresponded to Xu *et al.*'s [15] Behavioral Intention item worded as, "I intend to use mobile apps in the next 12 months." We consider the number of apps installed on the device as a good reference point for actual app use.

Total Dangerous Permissions Granted: The number of Dangerous Permissions granted to the apps installed on the device is captured by this data. Access to an Android user's location (fine as well as coarse) is considered a Dangerous Permission along with access to other sensitive information, such as calendar, call logs, camera, contacts, microphone, phone, sensors, SMS, storage, etc. [18]. Therefore, more Dangerous Permissions granted by the user should increase the likelihood that participants would grant the location-sharing permission to our app.

Location Ratio (percentage): We calculated Location Ratio as the number of installed apps with location permission granted divided by the total number of apps installed on the device. We expressed this ratio as a percentage. This variable serves to quantify past location-sharing behavior. If the assumption is that past behavior is the strongest predictor of future behavior, then this variable should have the highest predictive value.

Hybrid measures capturing users' perceptions of their past behavior

We combined our scraped behavioral variables with a self-report approach to create new hybrid measures that captured people's perceptions of their past behavior.

Location Comfort (percentage): This measure was calculated as the percentage of apps installed on the device to which the participant had granted location permission (i.e., behavioral data) and expressed comfort with the granted location access. This variable captures a person's reflection regarding past privacy-related behavior. The Theory of Planned Behavior [20] suggests that behavior is based on one's attitude toward that behavior. In this case, we asked study participants (Figure 1) to reflect explicitly on their past location-sharing behavior and self-report whether they felt comfortable about it (as "Yes" or "No"). In this way, the measure captures a subjective attitudinal component (i.e., comfort level) as well as an objective behavioral component (i.e., location permissions granted in the past) that contextualized participants' attitudes regarding their actual past location-sharing behavior.

Location Revoke (percentage): For the apps that participants were uncomfortable with accessing location, we further asked them to indicate if they will revoke this access. We calculated the Location Revoke measure as the percentage of apps for which a participant reported the desire to revoke location access because of being uncomfortable with sharing location (despite having granted location access to these apps earlier). Thus, the Location Revoke percentage variable captures the intention to revoke location access from apps with which location sharing was found to be uncomfortable. We used this measure as a proxy for understanding regrets about past privacy choices in anticipation of future changes in privacy behavior.

These hybrid measures quantify the outcome of reflecting on past behavior to make future privacy decisions. The reflective learning paradigm suggests that having a negative evaluation of past behavior should trigger changes in future behavior. Thus, we capture the extent to which people are unsatisfied with their current location-sharing choices and use this measure to predict their future behavior.

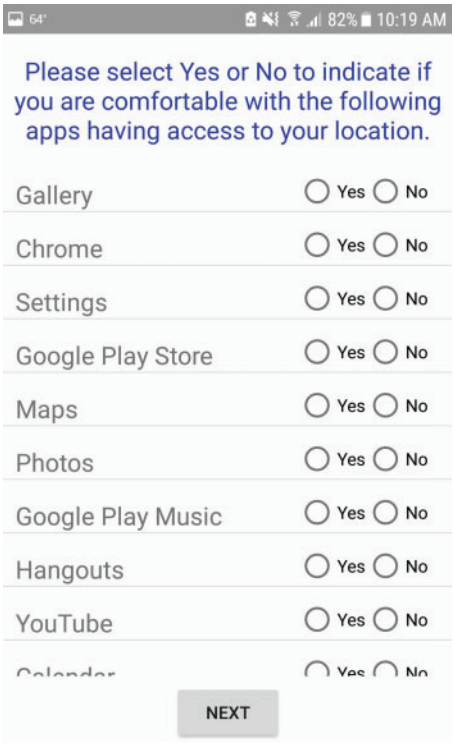

Figure 1: Location Comfort. Screenshot of the study app asking whether the participant was comfortable sharing location with the apps on the phone. This list was dynamically generated based on the apps installed on the participant's phone.

Methods

Our goal was to verify whether existing self-reported measures (RQ1A) and behavioral data (RQ1B) are suitable for predicting actual user behavior regarding location sharing. We further wanted to examine if these prediction models could be strengthened by creating a hybrid of these two classes of variables (RQ1C). To collect data on the variables relevant to tackling our research questions, we implemented an app for smartphones running the Android operating system. The following subsections describe the steps involved in the study deployment and participant recruitment. All study procedures were reviewed and approved by the Institutional Review Board (IRB) of the University of Central Florida (UCF).

Study design and app flow

As mentioned earlier, we were interested in three types of measures: 1) self-reported, 2) behavioral, and 3) hybrid. To collect these measures simultaneously and seamlessly, we implemented a smartphone app that incorporated an in-app questionnaire to collect the self-reported measures. While participants answered the questionnaire, the app collected information in the background regarding the apps installed on the device and the permissions granted to each of these apps. Then, we asked participants to reflect on this scraped behavioral information as a way to measure their perceptions around this objective data.

Figure 2 illustrates the various steps involved in the study. Android users recruited from Amazon Mechanical Turk who were interested in participating in the study were directed to a webpage that introduced the study and sought informed consent for participation. To avoid priming, we did not use the term 'privacy' anywhere within the study description. After reading the study description,

those who consented to participate in the study were provided with a randomly generated unique 'Consent ID' and directed to a link to install our study app from the Google Play app store. Upon installing and launching the app, each participant was first required to enter the Consent ID to verify completion of the informed consent procedures. All data collected by the app was transmitted to our database over a secure channel.

While the participant was answering the questionnaire, the app ran a background process to collect information on the apps installed on the device along with the permissions granted to each app. Note that the study description that sought informed consent explicitly disclosed the background data collection. Given the privacy-sensitive nature of this information, we minimized the extent of the collected permissions data by capturing information only for the permissions classified by Google as Dangerous Permissions. For each app on a participant's device, we collected the list of all granted Dangerous Permissions. We recorded whether each of the granted Dangerous Permissions was: (i) present in the respective app's manifest file but not explicitly requested from the device user, (ii) requested but denied by the device user, or (iii) requested and granted.

Upon completing the questionnaire, participants were presented with a list of all apps on their device that had been granted access to the location permission. As shown in Fig. 1, participants were asked to indicate whether they were comfortable sharing their location with each of the apps in the list. Once participants completed this step and chose to continue, our study app asked for access to the location permission using the standard permissions dialog of the Android operating system. Prior to presenting this decision, we explained, "We will ask you for your location when you press the next button. Please grant or deny us permission to store your location." We explicitly included the option for participants to deny access as we did not want to exert undue influence on their decision. Ostensibly, the study app's location request was made in order to enable us to collect participant location as one of the pieces of demographic information requested by the app at this point in the study. We used the participants' location-sharing decisions to record our dependent variable "Location Given." After recording participants' choices regarding providing location access to the study app, we requested demographic information and concluded the study.

Data analysis approach

The descriptive statistics for all our variables are provided in Table 1. To prepare our data for analysis, we first calculated Cronbach's alpha to assess the construct validity of our self-reported measures. Cronbach's alpha measures the internal consistency of a given construct [47], and a threshold of 0.7 is generally considered acceptable

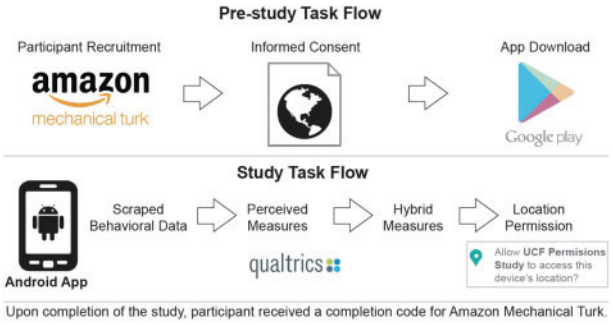

Figure 2: The flow of the various steps involved in the study.

[60]. All our self-reported constructs were above this threshold except Behavioral Intention ($\alpha = 0.65$) and Perceived Surveillance ($\alpha = 0.63$). Therefore, we examined the individual item measures in these scales more closely.

For Behavioral Intention, we conducted an Exploratory Factor Analysis (EFA) to understand why the internal consistency of our construct was below the suggested threshold. This exploratory analysis was appropriate given that we added two additional statements to the original pre-validated scale that were more specific to location-sharing. A principal component analysis with Varimax (orthogonal) rotation and Eigenvalues over one yielded two factors explaining a total of 78.7% of the cumulative variance across all scale items: (i) Behavioral Intent to Share Information with Apps and (ii) Behavioral Intention to Use Apps. Behavioral Intent to Share Information with Apps was composed of three statements (one from the Xu *et al.*'s original work [15] and our two statements specific to location sharing), and Behavioral Intention to Use Apps was composed of two statements from Xu *et al.*'s original work [15]. Therefore, we split this scale into two different types of Behavioral Intention (as described in Appendix 1).

For Perceived Surveillance, the belief that "the location of my mobile device is monitored at least part of the time" did not correlate well with the concern that mobile apps collect too much information or monitor one's activities. Dropping this item improved the internal consistency of Perceived Surveillance to $\alpha = 0.89$. After making these adjustments, the internal consistency of these self-reported scales were above the acceptable threshold for Cronbach's alphas (see Table 1). Once self-reported measures were confirmed for internal consistency, we created indices for these constructs by averaging across all scale items (see Appendix 1).

Next, we calculated the percentage of participants who reported that they would revoke location permission to an app for which they indicated in the previous step that they were uncomfortable with the app having their location. The screen we used for gathering this information was similar to the one in Fig. 1. We also measured the opposite, i.e., apps to which participants chose to allow location access when that app did not previously have such access. We did not differentiate between permissions for "coarse location" and "fine location" and included both when calculating Location Comfort percentages. The dialog requesting location sharing uses the same wording to request access without specifying which of the two types of location permissions is sought. Nonetheless, we calculated Location Comfort only for apps that were granted the "access fine location" permission and found that it made a negligible difference to our final model.

In preparation for our data analysis, we standardized all variables to their *z*-score. After standardizing our variables, we conducted binary logistic regression analyses to answer each of our high-level research questions. We used binary logistic regression because our dependent variable is dichotomous [62] (i.e., grant/deny location permission to our study app). First, we examined separate models for each of the three classes of variables to identify which self-reported measures (RQ1A), behavioral data (RQ1B), and self-reported in the context of past behavior measures (RQ1C) were significant in predicting our dependent variable. Finally, we performed a stepwise logistic regression that included the statistically significant variables from each model to combine them into a single model (RQ2).

Participant recruitment and sample characteristics

We recruited participants by posting the study as a Human Intelligence Task (HIT) on the Amazon Mechanical Turk crowd

Table 1: Descriptive statistics of independent variables and internal consistency for self-reported measures.

Variable	Variable type	Mean	Median	SD	Skewness	Kurtosis	Cronbach's alpha
Behavioral Intention to Share Information with Apps	Self-reported	3.76	4.00	0.991	−0.840	0.154	0.80
Behavioral Intention to Use Apps	Self-reported	4.41	4.50	0.749	−1.383	1.967	0.86
Perceived Surveillance	Self-reported	1.95	2.00	1.070	1.656	3.009	0.89
Perceived Intrusion	Self-reported	3.87	4.00	1.049	−1.284	1.993	0.90
Secondary Usage of Personal Information	Self-reported	4.19	4.33	0.876	−1.358	1.876	0.91
FYI Communication Style	Self-reported	2.51	2.33	1.068	0.308	−0.737	0.81
Power Usage	Self-reported	4.21	4.25	0.501	−0.644	0.358	0.79
Number of Installed Apps	Scraped behavioral	91.49	83.00	41.187	1.223	1.957	N/A
Total Dangerous Permissions Granted	Scraped behavioral	217.47	204.00	91.960	1.069	1.805	N/A
Location Ratio (percentage)	Scraped behavioral	26.23	25.65	8.678	0.332	−0.265	N/A
Location Comfort (percentage)	Hybrid	43.43	40.45	26.500	0.317	−0.875	N/A
Location Revoke (percentage)	Hybrid	24.52	13.84	27.487	1.101	0.073	N/A

work platform. To avoid the impact of cultural variance, we limited participation to U.S. adults (18 years of age or older). For adequate response quality, we restricted the HIT to workers who had HIT approval rates greater than 95% with at least 50 approved HITs. Since our study app could run only on the Android operating system, all participants were required to be users of Android devices. Upon completing the study, the app provided each participant with a randomly generated unique completion code to be entered on Amazon Mechanical Turk as the proof of completion of the study task. All participants who demonstrated successful completion of the study task by entering a valid completion code were compensated $1. While we did not record the time between participants consenting to participate in the study, installing the app, and completing the study tasks, pilot testing suggested that the duration of the study ranged from 15 to 25 minutes on average. Therefore, our statement of informed consent stated that "the study should take no longer than 30 minutes to complete." Participants were compensated even if they could not complete the study due to technical difficulties.

A total of 429 people accepted the HIT and completed the study tasks. After discarding the responses of those who failed the attention check embedded in the study, we ended up with valid data from 380 participants. We initially conducted an analysis with a subsample of 114 participants and published a preliminary paper on our results [63]. The present article is a follow-up to that initial study and includes the full set of 380 participants (over three times the sample size included in the initial report). The initial report [63] ($N = 114$) covered data collected between April 2018 and September 2018. We continued to recruit additional participants until November 2018. Given that there was no gap in data collection, we did not anticipate any systemic changes that would have significantly changed our results. T-tests on our model variables confirmed no statistical differences between the two samples collected until and after September 2018. Our high-level research questions and methods remained the same, but all statistical models were regenerated using the full data set, adding additional insight and nuance to our results. We highlight the differences between the two analyses in our discussion.

Our 380 participants included 195 (51.5%) males and 181 (47.8%) females. Two participants identified as "other," and one participant did not wish to specify. The participants came from various ethnic backgrounds including White/Caucasian (66.6%; $N = 253$), Black/African American (12.1%; $N = 46$), and Hispanic/Latino (6.8%; $N = 26$). Most of the participants (80.7%; $N = 306$) lived in urban or suburban areas (30.6% urban; $N = 116$ and 50.1% suburban; $N = 190$) with the remaining 19.3% ($N = 73$) coming from rural areas. Over half (55.7%; $N = 211$) of the participants reported completing at-least a 4-year college degree while almost one-third (32.7%; $N = 124$) reported completing no more than a 2-year college degree. More than half of the participants (53.9%; $N = 205$) were employed full time, while 14.7% ($N = 56$) were employed part-time, covering a diversity of occupations from Software Engineering to Food Management and earning a median income in the $40,000–$50,000 range. Another 6.3% participants ($N = 24$) were unemployed (e.g., looking for work, homemakers, or students).

Results

The following subsections describe the app use practices reported by our participants followed by the results of the analyses we carried out to answer our research questions.

Participants' mobile app use

When asked for location access by our study app (our dependent variable), 76.6% ($N = 291$) of the participants granted the permission. Our participants had an average of 91 apps installed on their devices with a minimum of 29 and a maximum of 272 and a standard deviation of 41. The top ten most common non-system apps were Messenger (62.4%), Facebook (60.3%), Duo (55.85%), Amazon Shopping (51.8%), Instagram (50.8%), Hangouts (45.3%), Netflix (39.5%), Docs (39.2%), Amazon Kindle (35.8%), and SmartThings (35.8%). On average, participants had granted 217 Dangerous Permissions to the various apps on their devices with an average of 24 of these apps having access to their location (fine or coarse). On average, participants were comfortable with about 43% of the location utilizing apps actually having location access. When asked if they would revoke location access for apps with which they were uncomfortable sharing location, 72.9% ($N = 277$) of the participants wished to do so for at least one such app.

In Appendix 2, we include a correlation matrix of the Pearson's bi-variate correlations between all study variables. Similar to Xu *et al.* [15], Behavioral Intention to Share Information with Apps was statistically significantly correlated with Behavioral Intention to Use Apps, Perceived Surveillance, Perceived Intrusion, and Secondary Usage of Personal Information. All signs of the coefficients were in expected directions. Further, Behavioral Intention to Share Information with Apps was statistically significantly and positively correlated with FYI Communication Style ($r = 0.212$), Power Usage ($r = 0.179$), Number of Installed Apps ($r = 0.246$), Total Dangerous Permissions Granted ($r = 1.99$), Location Ratio ($r = 0.114$), and Location Comfort ($r = 0.382$). It was statistically significantly and

Table 2: Binary logistic regression: Using self-reported measures to predict location sharing, Nagelkerke $R^2 = 16.51\%$.

Variable	Odds ratio	*P*	
Behavioral Intention to Share Information with Apps	**1.758**	**0.000**	***
Behavioral Intentionto Use Apps	1.119	0.415	
Perceived Surveillance	1.260	0.373	
Perceived Intrusion	1.103	0.652	
Secondary Usage of Personal Information	0.929	0.741	
FYI Communication Style	**1.327**	**0.044**	*
Power Usage	1.109	0.441	

$^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

negatively correlated with the percentage of apps tor which participants said they would revoke location sharing ($r = -0.358$). We noticed that the independent variables (i.e., Perceived Surveillance, Perceived Intrusion, and Secondary Usage of Personal Information) that predicted Behavioral Intention as a dependent variable were not significantly correlated with any of the scraped behavioral data (i.e., Number of Installed Apps, Total Dangerous Permissions, and Location Ratio). As expected, we observed a strong correlation between Number of Installed Apps and Total Dangerous Permissions ($r = 0.768$).

Binary logistic regression results

Self-reported Measures (RQ1A): Our first research question explored if self-reported measures can predict actual location-sharing behavior. The results of our logistic regression are shown in Table 2. Behavioral Intention to Share Information with Apps ($P < 0.000$, $e\beta = 1.758$) and FYI Communication Style ($P = 0.044$, $e\beta = 1.327$) were the only statistically significant predictors in this model. Behavioral Intention to Use Mobile Apps was not significant in our model. For each unit increase in Behavioral Intention to Share Information with Apps, participants were 1.76 times more likely to grant the location permission to our study app. For each unit increase in FYI Communication Style, participants were 1.33 times more likely to share their location with the study app. This model explained 16.5% of the variance in our dependent variable which was the highest variance explained across the three classes of independent variables by themselves.

Scraped Behavioral Measures (RQ1B): Next, we carried out a binary logistic regression using the scraped behavioral data (e.g., Number of Installed Apps, Total Dangerous Permissions Granted, and Location Ratio) as the independent variables. The results are shown in Table 3. Overall, we found these variables to be poor predictors of location sharing decisions with the model explaining only 3.2% of the variance. However, Number of Installed Apps ($P = 0.028$, $e\beta = 1.747$) and Location Ratio ($P = 0.028$, $e\beta = 1.479$) were statistically significant predictor variables. For each unit increase in Number of Installed Apps, the likelihood participants disclosed their location to our app increased 1.75 times. For each unit increase in Location Ratio, the likelihood participants disclosed their location to our app increased 1.48 times. Total Dangerous Permissions granted ($P = 0.092$, $e\beta = 0.995$) had no influence on location-sharing decisions.

Hybrid Measures (RQ1C): Overall, the binary logistica regression model with hybrid measures explained 12.4% of the variance in location-sharing decisions, outperforming scraped behavioral data considered by itself. The model (Table 4) shows that Location Comfort percentage was significant ($P < 0.000$, $e\beta = 1.829$), but

Table 3: Binary logistic regression: Using scraped behavioral measures to predict location sharing, Nagelkerke $R^2 = 3.2\%$.

Variable	Odds ratio	*P*	
Number of Installed Apps	**1.747**	**0.028**	*
Total Dangerous Permissions Granted	0.995	0.092	
Location Ratio	**1.479**	**0.028**	*

$^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

Location Revoke percentage ($P = 0.089$, $e\beta = 0.800$) was not. For each unit increase in Location Comfort percenta, participants were 1.83 times more likely to share their location with our study app.

Combined Model (RQ2): Next, we combined the results of the above three separate models to derive the best model given our data. RQ2 investigated what combination of factors across all three types of predictor variables best predicts app location-sharing behavior. In order to answer this research question, we combined all significant variables from the previous regression models (so as not to inflate our R^2 with variables that were not significant) to achieve the best model for predicting our dependent variable. We summarize these models and the R^2 change in Table 5. Adding the hybrid measure for Location Comfort percentage to Behavioral Intention to Share Information with Apps and the FYI Communication Style explained statistically significantly more variance in location-sharing behavior than the generic self-reported measures alone. When the Number of Installed Apps and Location Ratio were combined with the self-reported measures (Step 2), they became non-significant in the model. In our final model (Step 3), Location Comfort remained statistically significant.

According to the final model, each unit increase in Behavioral Intention to Share Information with Apps led to the likelihood of sharing location with our app increasing 1.63 times. For each unit increase in FYI Communication Style, the likelihood increased 1.36 times, and for Location Comfort, it increased 1.63 times. Based on these odds ratios, Location Comfort had an effect of similar strength on our dependent variable as that of Behavioral Intention to Share Information with Apps.

The change in explanatory power between the first model with only self-reported measures (Step 1) and final model with all three variable types (Step 3) was statistically significant. However, the difference when adding the scraped behavioral variables (Step 2) was not significant. This result suggests that hybrid measures that capture users' perceptions of their past behavior added value to the overall model and are better predictors of future behavior than past behavior considered by itself.

Discussion

For RQ1A, we found that self-reported measures could predict actual privacy behavior (i.e., location sharing) fairly well. Our model showed that Behavioral Intention to Share Information with Apps and FYI Communication Style were the most important factors in predicting whether the participants shared their location with our study app. For RQ1B, scraped behavioral data proved to be the worst at predicting privacy behavior. Only Number of Installed Apps was statistically significant, and the R^2 value of this model was very low. For RQ1C, our new hybrid measures showed notable predictive power for explaining location-sharing behavior. Location Comfort percentage was found to be a statistically significant predictor. Finally, the combined model of self-reported and hybrid

measures was the best predictor model of location sharing decisions. Behavioral Intention to Share Information with Apps, FYI Communication Style, and Location Comfort were found to be the statistically significant predictors in this model (RQ2). This result suggests that hybrid variables that capture users' perceptions about their past privacy choices can be used to augment traditional survey measures and scraped behavioral data to produce stronger predictive models of user behavior. We discuss the implications of these findings in detail below.

Implications for research on location-sharing and mobile privacy

In our review of the literature, most mobile privacy research we encountered was strictly divided between the social sciences or computational sciences, with few studies at the intersection of the two disciplines. Yet, we found that the best model for predicting smartphone users' app location-sharing behavior was a hybrid of the two. Below, we reflect on the implications of our results for these different privacy research communities individually and suggest a path forward that leverages the strengths of both approaches.

Social Science Privacy Research: Behavioral Intention is cited in the social sciences as the strongest predictor of actual behavior [20], and we confirmed that this holds true to some extent. Behavioral Intention to Share Information with Apps was the strongest predictor variable in our model, but this was only after we differentiated it from Behavioral Intention to Use Apps. Behavioral Intention to Use Apps was conceptually different to our participants than the intent to share personal information with these apps, and it was not a significant predictor of their location-sharing behavior. Overall, Behavioral Intention to Use Apps (M = 4.41, SD = 0.749) was higher than the Behavioral Intention to Share Information with Apps (M = 3.76, SD = 0.991). Similarly, all our participants were willing to install our study app but only about three quarters of them (76.6%; $N = 291$) granted our app access to their location. Thus, one contribution of our work is the refinement of Xu *et al.*'s [15] original scale for Behavioral Intention in the context of mobile privacy. The intention to use apps versus share information with them is discernably different, and future research should take this difference into consideration.

Similar to Behavioral Intention to Use Apps, power user status was high among our participants (M = 4.21, SD = 0.501); therefore, the lack of statistical significance in these variables may have been due to the proclivity of our participants to use mobile apps which does not (and arguably should not) equate to a propensity for sharing sensitive information, such as location, with these apps. Additionally, perceptions of mobile surveillance may have shifted over time. Participants' beliefs that their mobile phone monitored their location was not correlated with their concern that mobile apps collected too much information. This may be because modern-day Android users see location-based app services as commonplace. According to a 2016 Pew research study, 90% of smartphone users use their phone to get directions and other location-based services. This has increased from the 74% reported in 2013 [61]. Therefore, an important implication of our findings is that self-reported measures should be contextualized to the actual behavior the research intends to predict, particularly in the case of behavioral intent. In our case, contextualizing behavioral intent to that of sharing personal information with apps, specifically location, improved the predictive power of our model.

We make two additional important points regarding the use of self-reported constructs in survey-based social science research. First, we demonstrated that pre-validated measures from the literature can become antiquated with time because the context and meaning of the scale items can change. For example, the item about location tracking within the Perceived Surveillance construct likely invoked very different responses in 2018 than they did in 2004. Unlike mobile phone users over a decade ago, most modern Android

Table 4: Binary logistic regression: Using self-reported in the context of past behavior measures to predict location sharing, Nagelkerke $R^2 = 12.4\%$.

Variable	Odds ratio	*P*	
Location Comfort	**1.829**	**0.000**	***
Location Revoke	0.800	0.089	

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

Table 5: Stepwise binary logistic regression for comparing models.

Variable	Odds ratio	*P*		Nagelkerke R^2
Step 1				
Behavioral Intention to Share Information with Apps	**1.942**	**0.000**	***	15.22%
FYI Communication Style	**1.357**	**0.026**	*	
Step 2				
Behavioral Intention to Share iInformation with Apps	**1.899**	**0.000**	***	15.71%
FYI Communication Style	**1.375**	**0.022**	*	
Number of Installed Apps	1.022	0.876		
Location Ratio	1.158	0.256		
χ^2(Step1, Step2)=1.377, degrees of freedom=1, *P*=0.502				
Step 3				
Behavioral Intention to Share Information with Apps	**1.629**	**0.000**	***	19.49%
FYI Communication Style	**1.357**	**0.030**	*	
Number of Installed Apps	1.005	0.969		
Location Ratio	1.132	0.337		
Location Comfort Percentage	**1.633**	**0.002**	**	
χ^2(Step2, Step3)=10.779, degrees of freedom=1, *P*=0.001**				

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

users expect at least some of the apps on their phone to track location as standard practice. Second, we would not have detected this type of conceptual difference in the scale items using a smaller sample size. In our initial analysis [63] that included only 114 Android users, the original scale for Behavioral Intention from Xu *et al*. [15] demonstrated adequate internal consistency ($\alpha = 0.81$). It was only after we tripled the sample size ($N = 380$) that we were able to tease out the conceptual differences in these itemsthat clearly impacted our overall results. In our initial analysis [63], Behavioral Intention was not a significant predictor of location-sharing behavior, to our surprise. However, this was likely because the intention to use apps is not predictive of information sharing and created entropy in the Behavioral Intention construct. Therefore, we urge social scientists to continually interrogate the construct validity of the self-reported measures used in their studies becauseit can significantly impact the overall results, as it did ours.

We found that the FYI Communication Style influenced participants' location-sharing behavior, consistent with Page *et al*.'s [8] earlier results. FYI Communication Style pertains to how one communicates location information with others. The FYI Communication Style being a significant predictor of location-sharing behavior as well suggests that a user's preferred communication style is an important factor to consider when predicting whether the user will grant an app the permission to access location. Designers should consider personal preferences for the convenience of letting others (apps in this case) decide when location access is needed versus wanting a more explicit hands-on approach to disclosing location on a case-by-case basis. Further, we found that the FYI Communication Style personal trait was a stronger predictor of location-sharing behavior than any of the antecedents of Behavioral Intention (i.e., Perceived Surveillance, Perceived Intrusion, and Secondary Usage of Personal Information) in Xu *et al*.'s earlier work based on MUIPC [10] and the Theory of Planned Behavior [20]. This suggests that personal traits may play a bigger role than self-reported privacy concerns when deciding whether to grant location access. While Behavioral Intention to Use Apps may be useful for predicting some behaviors, such as technology adoption, privacy researchers should consider using more contextualized measures for the specific type of information disclosure being studied, similar to the FYI Communication Style personal trait that was created based on empirical work specifically on mobile location privacy sharing [8].

While Behavioral Intention to Share Information with Apps was a strong predictor of past privacy behavior, Behavioral Intention to Use Apps and the antecedent variables that measured privacy concerns did not add any value to our models. Regardless of why these self-reported measures did not correlate with any of the participants' scraped behavioral data or actual location-sharing behavior, our results call into question the common practice of studying privacy concerns as a proxy for understanding actual privacy behavior. Because privacy behaviors are often paradoxical, and possibly more nuanced, than other technology-related behaviors, such as technology adoption [64, 65], it is possible that a generalized Behavioral Intention construct [10] and the Theory of Planned Behavior [20] may not be the best proxies or approaches for predicting actual privacy behavior. Similarly, Power Usage was not a significant predictor of participants' location-sharing behavior. Upon further reflection, this lack of an effect may be because these constructs focus on using technology in general, whereas Behavioral Intention to Share Information with Apps and FYI Communication Style are more directly connected to the concept of location sharing in particular. These results indicate that location-specific self-reported constructs should be used to predict users' location-sharing behavior, rather than more general measures about privacy concerns or mobile app use. Finally, social science researchers should consider moving beyond using Behavioral Intention as a dependent variable and treat it as an independent variable that predicts actual behavior.

Computational Privacy Research: Scraped behavioral data was not a good predictor of location disclosure at all which suggests that users' future location disclosure behavior is not necessarily tied to their past behavior. The Number of Installed Apps and the number of Dangerous Permissions granted may be too broad and thus not tied to attitudes about location information. However, we were surprised that past location-sharing choices (i.e., Location Ratio) were not a significant predictor of future location-sharing behavior. While Ghosh *et al*. [43] found that device metadata, such as call duration and ignored calls, could predict self-reported privacy concerns, our findings suggest that phone metadata might not be the best predictor of actual privacy behavior such as location sharing. In fact, past choices may not even reflect desired behavior; permissions may have been granted as a condition of using the app, without much thought or understanding of the implications of the decision. As a result, currently granted permissions do not seem to predict the desired behavior when the user is explicitly asked to grant location permissions. Therefore, we caution computational researchers and system designers against using past privacy behavior as the sole proxy for determining users' future privacy preferences when creating predictive models, designing "intelligent" defaults, and recommending privacy choices.

The Importance of Self-Reflecting on Past Behavior: Our key novel findings were that people's reflection on their past behavior (i.e., their comfort with their past location-sharing decisions) was a predictor of their future behavior (i.e., whether they granted location access to our study app). Incorporating self-reflection on past behavior significantly improved our models beyond using only non-contextual self-report (i.e., generic self-reported measures) or past actions (i.e., scraped behavioral data) as predictor variables. Of the self-reported measures where participants reflected on their past location-sharing choices, Location Comfort was the most influential in improving the prediction of location-sharing behavior, more so than the Location Revoke measure. This suggests that users' comfort with their past location permissions translates directly to their future decisions about sharing their location. Therefore, future research needs to look beyond raising users' awareness of their privacy behavior and try to help them feel comfortable about their mobile privacy settings by reflecting on past choices.

Overall, the results of our study suggest that combining the social science and computer science approaches can yield stronger predictive models. Our work encourages future privacy research to identify and measure relevant user perceptions about past privacy choices in the context of the privacy decision currently being made. In our case, we identified Location Comfort percentage as the hybrid (i.e., self-reported in the context of past behavior) variable strongly correlated with location-sharing decisions. This suggests that users' self-evaulations of their past behavior are a better predictor of future behavior than the actual past behavior itself.

Implications for design

Prior research has shown that achieving the right "privacy fit" can lead to higher user engagement with the service and help users feel more socially connected with others [4]. Our work shows that user perceptions of their past behavior (in our case, location-sharing decisions) greatly improves on using just attitudes about future

disclosure to predict future behavior. Therefore, a combination of attitudinal measures asked in the context of one's actual past behavior is more useful for understanding attitudes that lead to concrete action. In fact, this technique mirrors the "reflective learning" [49] approach that has been shown to produce positive learning outcomes [48]. Namely, by reflecting on one's past choices, one can become more aware of one's actions and make better future choicesthat might be incongruous with one's past choices. This capability to "learn reflectively" could be supported in the design of apps as a context-aware feature that periodically reminds users of their past decisions and gives them the opportunity to reflect and change their decisions based on context changes or bad experiences with previous decisions. Such designs could support a more dynamic conception of privacy that matches what users want, as opposed to what they think they want or what they did in the past.

Limitations and future research

Our study has several limitations that can inform future research. First, limiting participation only to those from the United States constrains our ability to generalize our results to other populations because privacy decisions and experiences can be shaped by the cultural environment. Since we limited participation to adults of ages 18 and above, the applicability of our results to younger populations needs to be verified. Further, it may be useful to verify whether our results generalize to those who use devices with other operating systems, such as Apple's iOS.

Through our analyses, we refined existing pre-validated measures for mobile privacy (i.e., Behavioral Intention and Perceived Surveillance). While the contributions of our research are primarily empirical, our exploratory results based on the internal consistency of our measures warrant future research that uses more confirmatory approaches to validate the psychometric properties of our revised scales. Another promising area of future research would be examining additional factors beyond the ones incorporated in this study. For instance, it would be useful to study whether location-sharing decisions are affected by external factors, such as news about privacy or data breaches, in addition to users' reflections on their past privacy behavior.

We recruited Amazon Mechanical Turk workers for the purpose of this research which presents both limitations and ethical considerations. Kang *et al.*'s work [66] suggests that Amazon Mechanical Turk workers have unique privacy profiles compared to average users. Moreover, individuals recruited from Amazon Mechanical Turk are likely to be more technically savvy than the general population. As a result, the use of a participant sample recruited from Amazon Mechanical Turk may constrain the generalizability of our results. While participation in our study was voluntary, we acknowledge that the compensation for our study was below minimum wage. Two assistant professors, who were co-authors of this work, paid participant support costs using limited funds provided by their universities. Unfortunately, since this was preliminary work for future grant proposals (acknowledged at the end of this article), we did not have the resources to pay participants an amount equivalent to the minimum wage. Busse *et al.* [67] addressed this concern of fair payment by retroactively paying Amazon Mechanical Turk workers an additional bonus. Had we been aware of this mechanism, we could have budgeted for the bonus in our grant proposals. Our experience could help other researchers consider appropriate payment structures when recruiting and compensating participants on Amazon Mechanical Turk.

Future research should consider replicating our study with samples drawn from other populations that are more diverse in terms of ages, cultures, and technical abilities and identify more context-specific hybrid measures that ask users to self-reflect on their past behavior. Such contextualization will not only help improve predictive models of behavior, but also enhance the user experience by customizing our research to the unique experiences of the participants. An interesting future direction could be to track changes in location-sharing comfort for specific apps over the duration of time they are installed on the phone. Such tracking could help app makers identify location-sharing behavior trends and take corrective action.

Conclusion

The rapid growth of mobile devices has led to the boundaries of privacy being tested in new ways. It is therefore crucial to understand and be able to predict user behavior in order to design experiences which respect user expectations of privacy. We contribute to the field of mobile privacy by shedding light on the kinds of self-reported measures that can explain user location-sharing behavior. We show that scraped behavioral data might not be the best indicator of future user behavior. However, augmenting self-reported measures with users' perceptions of their past behavior can help strengthen prediction models.

Acknowledgements

We would like to thank those who contributed to this work. Heather Lipford and Bart Knijnenburg gave input on our study design. Malak Eihab Aly helped develop and test the app we used to conduct the study. Abhiditya Jha assisted with data collection, statistical modeling, and paper formatting. Finally, we value the participants who took the time to participate our study.

Funding

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.

Appendix 1

Self-reported measures

Behavioral intention to share information with apps

Item 1 was adapted from Xu *et al.* [15], and we created items 2 and 3 given the specific context of our study. These items were measured on a 5-point Likert scale from 1 – Agree Strongly, 2 – Agree Somewhat, 3 – Neutral, 4 – Disagree Somewhat, 5 – Disagree Strongly.

Items:

1. I am likely to disclose my personal information to use mobile apps in the next 3 months.
2. I am likely to grant permission to share my location with my existing mobile apps in the next 3 months.
3. I am likely to grant permission to share my location with new mobile apps in the next 3 months.

Behavioral intention to use apps

Both items were adapted from Xu *et al*. [15], these items were measured on a 5-point Likert scale from 1 – Agree Strongly, 2 – Agree Somewhat, 3 – Neutral, 4 – Disagree Somewhat, 5 – Disagree Strongly.

Items:

1. I predict I will use new mobile apps in the next 3 months.
2. I intend to use mobile apps in the next 3 months.

Perceived surveillance

Taken from Xu *et al*. [10], these items were measured on a 5-point Likert scale from 1 – Agree Strongly, 2 – Agree Somewhat, 3 – Neutral, 4 – Disagree Somewhat, 5 – Disagree Strongly.

Item removed:

1. I believe that the location of my mobile device is monitored at least part of the time.

Items included:

2. I am concerned that mobile apps are collecting too much information about me.
3. I am concerned that mobile apps may monitor my activities on my mobile device.

Perceived intrusion

Taken from Xu *et al*. [16], these items were measured on a 5-point Likert scale from 1 – Agree Strongly, 2 – Agree Somewhat, 3 – Neutral, 4 – Disagree Somewhat, 5 – Disagree Strongly.

Items:

1. I feel that as a result of my using mobile apps, others know about me more than I am comfortable with.
2. 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.
3. I feel that as a result of using mobile apps, information about me is out there that, if used, will invade my privacy.

Secondary use of personal information

Taken from Smith *et al*. [17], these items were measured on a 5-point Likert scale from 1 – Agree Strongly, 2 – Agree Somewhat, 3 – Neutral, 4 – Disagree Somewhat, 5 – Disagree Strongly.

Items:

1. I am concerned that mobile apps may use my personal information for other purposes without notifying me or getting my authorization.
2. When I give personal information to use mobile apps, I am concerned that apps may use my information for other purposes.
3. I am concerned that mobile apps may share my personal information with other entities without getting my authorization.

FYI communication style 'about myself'

Taken from Page et al. [8], these items were measured on a 5-point Likert scale from 1 - Agree Strongly, 2 – Agree Somewhat, 3 - Neutral, 4 - Disagree Somewhat, 5 – Disagree Strongly.

Items:

1. I want others to know what I am up to without my having to bother to tell them.
2. Others should be able to find out about me when they feel they need to.
3. I would prefer to share about myself with everyone in case anyone wants to know.

Power usage

Taken from Marathe et al. [7], these items were measured on a 5-point Likert scale from 1 - Agree Strongly, 2 – Agree Somewhat, 3 - Neutral, 4 - Disagree Somewhat, 5 – Disagree Strongly.

Items:

1. I think most technological gadgets are complicated to use.
2. I make good use of most of the features available in any technological device.
3. I have to have the latest available upgrades of technological devices that I use.
4. Use of information technology has almost replaced my use of paper.
5. I love exploring all the features that any technological gadget has to offer.
6. I often find myself using many technological devices simultaneously.
7. I prefer to ask friends how to use any new technological gadget instead of trying to figure it out myself.
8. Using any technological device comes easy to me.
9. I feel like information technology is a part of my daily life.
10. Using information technology gives me greater control over my work environment.
11. Using information technology makes it easier to do my work.
12. I would feel lost without information technology.

Appendix 2
Correlations matrix of all independent variables

$N = 380$ in all cases	Behavioral intention to share information with apps	Behavioral intention to use apps	Perceived surveillance	Perceived intrusion	Secondary usage of personal information	FYI communication style	Power usage	Number of installed apps	Total dangerous permissions granted	Location ratio (percentage)	Location comfort (percentage)	Location revoke (percentage)
Behavioral Intention to Share Information with Apps	1	0.408***	−0.205***	−0.221***	−0.198***	0.212***	0.179***	0.246***	0.199***	0 .114*	0.382***	−0.358***
Behavioral Intention to Use Apps		1	.002	−0.009	−0.018	0.048	0.376***	0.150**	0.138**	0.094	0.197***	−0.161**
Perceived Surveillance			1	0.734***	0.741***	−0.189***	0.063	0.021	0.006	−0.024	−0.261***	0.261***
Perceived Intrusion				1	0.694***	−0.177***	0.023	−0.072	−0.066	−0.023	−0.249***	0.206***
Secondary Usage of Personal Information					1	−0.157**	0.055	−0.032	−0.022	−0.047	−0.265***	0.279***
FYI Communication Style						1	0.115*	0.055	0.079	−0.021	0.139**	−0.111*
Power Usage							1	0.227***	0.205***	0.011	0.087	−0.056
Number of Installed Apps								1	0.768***	0.101*	0.131*	−0.209***
Total Dangerous Permissions Granted									1	0.517***	0.078	−0.167**
Location Ratio (percentage)										1	0.078	−0.144**
Location Comfort (percentage)											1	−0.512***
Location Revoke (percentage)												1

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

References

1. Global social media research summary 2019. Smart Insights. https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/ (24 July 2019, date last accessed)
2. App Monetization: Store, Ads to Deliver $189B by 2020. https://www.appannie.com/en/insights/market-data/app-monetization-report-2016/ (24 July 2019, date last accessed)
3. Olmstead K, Atkinson M. Apps Permissions in the Google Play Store. *Pew Research Center*. http://www.pewinternet.org/2015/11/10/apps-permissions-in-the-google-play-store/, 2015.
4. Wisniewski P, Isam A, Knijnenburg BP *et al.* Give social network users the privacy they want. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pp. 1427–41. New York, NY, USA: ACM, 2015.
5. Barth S, de Jong MDT. The privacy paradox: investigating discrepancies between expressed privacy concerns and actual online behavior – a systematic literature review. *Telematics and Informatics* 2017;**34**:1038–58.
6. Norberg PA, Horne DR, Horne DA. The privacy paradox: personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs* 2007;**41**:100–26.
7. Marathe S, Sundar SS, Nije Bijvank M *et al. Who are these power users anyway? Building a psychological profile*. In *Proceedings 57th Annual Conference of the International Communication Association* 2007.
8. Page X, Knijnenburg BP, Kobsa A. FYI: Communication style preferences underlie differences in location-sharing adoption and usage. In: *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 153–62. New York, NY, USA: ACM, 2013.
9. Sundar SS, Marathe SS. Personalization versus customization: the importance of agency, privacy, and power usage. *Human Communication Research* 2010;**36**:298–322.
10. Xu H, Gupta S, Rosson MB *et al.* Measuring mobile users' concerns for information privacy. In: *Proceedings of the Thirty Third International Conference on Information Systems,* 2012.
11. Almuhimedi H, Schaub F, Sadeh N *et al.* 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*, pp. 787–796 New York, NY, USA: ACM, 2015.
12. Amos B, Turner H, White J. Applying machine learning classifiers to dynamic Android malware detection at scale. In: *9th International Wireless Communications and Mobile Computing Conference*, Sardinia, Italy, pp. 1666–71, 2013.
13. Li Y, Guo Y, Chen X. PERUIM: Understanding mobile application privacy with permission-UI mapping. In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 682–93. New York, NY, USA: ACM, 2016.
14. Wisniewski P, Islam A, Lipford HR *et al.* Framing and measuring multidimensional interpersonal privacy preferences of social networking site users. *Communications of the Association for Information Systems* 2016;**38**:235–58.
15. Xu H, Teo H-H. Alleviating consumers' privacy concerns in location-based services: a psychological control perspective. In: *Proceedings of the Twenty-Fifth International Conference on Information Systems*. Washington, DC, USA, pp. 793–806, 2004.
16. Xu H, Dinev T, Smith HJ, *et al.* Examining the formation of individual's privacy concerns: toward an integrative view. In: *Proceedings of the Twenty-Ninth International Conference on Information Systems*, Paris, France, pp. 1–16, 2008.
17. Smith HJ, Milberg SJ, Burke SJ. Information privacy: measuring individuals' concerns about organizational practices. *MIS Quarterly* 1996;**20**:167–96.
18. Google Inc. Permissions Overview. https://developer.android.com/guide/topics/permissions/overview (2 January 2019, date last accessed)
19. Preibusch S. Guide to measuring privacy concern: review of survey and observational instruments. *International Journal of Human-Computer Studies* 2013;**71**:1133–43.
20. Ajzen I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 1991;**50**:179–211.
21. Malhotra NK, Kim SS, Agarwal J. Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. *Information Systems Research* 2004;**15**:336–55.
22. Buchanan T, Paine C, Joinson AN *et al.* Development of measures of online privacy concern and protection for use on the Internet. *Journal of the American Society for Information Science and Technology* 2007;**58**:157–65.
23. Keith MJ, Babb JS, Lowry PB. A longitudinal study of information privacy on mobile devices. In: *47th Hawaii International Conference on System Sciences,* Big Island, Hawaii, USA, pp. 3149–58, 2014.
24. Sharma S, Crossler RE. Disclosing too much? Situational factors affecting information disclosure in social commerce environment. *Electronic Commerce Research and Applications* 2014;**13**:305–19.
25. Guha S, Wicker SB. Spatial subterfuge: an experience sampling study to predict deceptive location disclosures. In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 1131–35. New York, NY, USA: ACM, 2015.
26. Page X, Kobsa A, Knijnenburg BP. Don't disturb my circles! Boundary preservation is at the center of location-sharing concerns. In: *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, Dublin, Ireland, pp. 266–73, 2012.
27. Barkhuus L, Dey A. Location-based services for mobile telephony: a study of users' privacy concerns. In: *Proceedings of Interact 2003*, Zurich, Switzerland, pp. 709–12, 2003.
28. Liu B, Lin J, Sadeh N. Reconciling mobile app privacy and usability on smartphones: could user privacy profiles help? In: *Proceedings of the 23rd International Conference on World Wide Web*, pp. 201–12. New York, NY, USA: ACM, 2014.
29. Chitkara S, Gothoskar N, Harish S *et al.* Does this app really need my location?: Context-aware privacy management for smartphones. *Proceedings of the ACM on Interactive Mobile, Wearable and Ubiquitous Technologies* 2017;**1**:1–42:22.
30. Srivastava G, Chitkara S, Ku K *et al.* PrivacyProxy: leveraging crowdsourcing and in situ traffic analysis to detect and mitigate information leakage. *CoRR* 2017; abs/1708.06384.
31. Fu H, Yang Y, Shingte N *et al.* A field study of run-time location access disclosures on android smartphones. In: *Proceedings of the Workshop on Usable Security*, San Diego, CA, USA, 2014;**14**.
32. Liu B, Andersen MS, Schaub F *et al.* Follow my recommendations: a personalized privacy assistant for mobile app permissions. In: *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, pp. 27–41. Denver, CO: USENIX Association, 2016.
33. Fawaz K, Shin KG. Location privacy protection for smartphone users. In: *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pp. 239–50. New York, NY, USA: ACM, 2014.
34. Wijesekera P, Reardon J, Reyes I *et al.* Contextualizing privacy decisions for better prediction (and protection). In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 268:1–268:13. New York, NY, USA: ACM, 2018.
35. Olejnik K, Dacosta I, Machado JS *et al.* SmarPer: context-aware and automatic runtime-permissions for mobile devices. In: *2017 IEEE Symposium on Security and Privacy (SP)*, San Jose, CA, USA, pp. 1058–76. 2017.
36. Wang W, Wang X, Feng D *et al.* Exploring permission-induced risk in Android applications for malicious application detection. *IEEE Transactions on Information Forensics and Security* 2014;**9**:1869–82.
37. Moonsamy V, Rong J, Liu S. Mining permission patterns for contrasting clean and malicious android applications. *Future Generation Computer Systems* 2014;**36**:122–32.
38. Mylonas A, Theoharidou M, Gritzalis D, Assessing privacy risks in Android: a user-centric approach. In: Bauer T, Großmann, J, Seehusen F *et al.* (eds). *Risk Assessment and Risk-Driven Testing*. Cham: Springer International Publishing, 2014, 21–37.
39. Jeon J, Micinski KK, Vaughan JA *et al. Dr. Android and Mr. Hide: Fine-Grained Security Policies on Unmodified Android*, 2011.

40. Wisniewski PJ, Knijnenburg BP, Lipford HR. Making privacy personal: profiling social network users to inform privacy education and nudging. *International Journal of Human-Computer Studies* 2017; **98**:95–108.
41. Lin J, Amini S, Hong JI *et al.* Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 501–10. New York, NY, USA: ACM, 2012.
42. Lin J, Liu B, Sadeh N *et al.* Modeling users' mobile app privacy preferences: restoring usability in a sea of permission settings. In: *Proceedings of the 10th Symposium on Usable Privacy and Security*, pp. 199–212. Menlo Park, CA: USENIX Association, 2014.
43. Ghosh I, Singh VK, Predicting privacy attitudes using phone metadata. In: Xu KS, Reitter D, Lee D *et al.* (eds). *Social, Cultural, and Behavioral Modeling*, Vol **9708**. Cham: Springer International Publishing, 2016, 51–60.
44. Seneviratne S, Seneviratne A, Mohapatra P *et al.* Predicting user traits from a snapshot of apps installed on a smartphone. *SIGMOBILE Mobile Computing and Communication Reviews* 2014;**18**:1–8.
45. Seneviratne S, Seneviratne A, Mohapatra P *et al.* Your installed apps reveal your gender and more*! SIGMOBILE Mobile Computing and Communication Reviews* 2015;**18**:55–61.
46. Malmi E, Weber I. You are what apps you use: demographic prediction based on user's apps. In: *Tenth International AAAI Conference on Web and Social Media*. Cologne, Germany, 2016.
47. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychological Bulletin* 1955; **52**:281–302.
48. Loo R, Thorpe K. Using reflective learning journals to improve individual and team performance. *Team Performance Management: An International Journal* 2002;**8**:134–39.
49. Boyd EM, Fales AW. Reflective learning: key to learning from experience. *Journal of Humanistic Psychology* 1983; **23**:99–117.
50. Raento M, Oulasvirta A. Designing for privacy and self-presentation in social awareness. *Personal and Ubiquitous Computing* 2008;**12**:527–42.
51. Consolvo S, Smith IE, Matthews T *et al.* Location disclosure to social relations: why, when, & what people want to share. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 81–90. New York, NY, USA: ACM, 2005.
52. Iachello G, Smith I, Consolvo S *et al.* Control, deception, and communication: evaluating the deployment of a location-enhanced messaging service. In: Beigl M, Intille S, Rekimoto J *et al.* (eds). *UbiComp 2005: Ubiquitous Computing*. Berlin; Heidelberg: Springer, 2005, 213–31.
53. Brown B, Taylor AS, Izadi S *et al.* Locating family values: a field trial of the whereabouts clock. In: Krumm J, Abowd GD, Seneviratne A *et al.* (eds). *UbiComp 2007: Ubiquitous Computing*. Berlin; Heidelberg: Springer, 2007, 354–71.
54. Shih F, Liccardi I, Weitzner D. Privacy tipping points in smartphones privacy preferences. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 807–16. New York, NY, USA: ACM, 2015,.
55. Wiese J, Kelley PG, Cranor LF *et al.* Are you close with me? Are you nearby?: Investigating social groups, closeness, and willingness to share. In: *Proceedings of the 13th International Conference on Ubiquitous Computing*, pp. 197–206. New York, NY, USA: ACM, 2011.
56. Boesen J, Rode JA, Mancini C. The domestic panopticon: location tracking in families. In: *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, 65–74. New York, NY, USA: ACM, 2010.
57. Kennedy-Lightsey CD, Martin MM, Thompson M *et al.* Communication privacy management theory: exploring coordination and ownership between friends. *Communication Quarterly* 2012;**60**:665–80.
58. Solove DJ. A taxonomy of privacy. *University of Pennsylvania Law Review* 2006;**154**:477–560. Vol
59. Kang H, Shin W. Do smartphone power users protect mobile privacy better than nonpower users? Exploring power usage as a factor in mobile privacy protection and disclosure. *Cyberpsychology, Behavior, and Social Networking* 2016;**19**:179–85.
60. Cho E, Kim S. Cronbach's coefficient alpha: well known but poorly understood. *Organizational Research Methods* 2015;**18**:207–30.
61. NW 1615 L. St, Suite 800Washington, Inquiries D 20036USA202-419-4300 | M-857-8562 | F-419-4372 | MONICA A. More Americans using smartphones for getting directions, streaming TV. *Pew Research Center* 2016, https://www.pewresearch.org/fact-tank/2016/01/29/us-smartphone-use/.
62. Kleinbaum DG, Klein M. *Logistic Regression: A Self-Learning Text*. Statistics for Biology and Health. 3rd edn. New York: Springer Science & Business Media, 2010.
63. Safi, M.I., Jha, A., Aly, M.E., Page, X., Patil, S., and Wisniewski, P. Will They Share? Predicting Location Sharing Behaviors of Smartphone Users through Self-Reflection on Past Privacy Behaviors, in the Proceedings of the 2019 NDSS Workshop on Usable Security and Privacy (USEC 2019), San Diego, CA. (2019), 10.14722/usec.2019.23014.
64. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences* 2008; **39**:273–315.
65. Venkatesh V, Morris MG, Davis GB *et al.* User acceptance of information technology: toward a unified view. *MIS Quarterly* 2003;**27**: 425–78.
66. Kang R, Brown S, Dabbish L *et al.* Privacy attitudes of mechanical Turk workers and the U.S. Public. In: *Proceedings of the 10th Symposium on Usable Privacy and Security*, pp. 37–49. Menlo Park, CA: USENIX Association, 2014.
67. Busse K, Schafer J, Smith M. Replication: no one can hack my mind revisiting a study on expert and non-expert security practices and advice. In: *Proceedings of the 15th Symposium on Usable Privacy and Security*, pp. 117–36. Santa Clara, CA: USENIX Association, 2019.