Too Much Too Late: Influence of risk communication on Android App installations – UNDER REVIEW

Prashanth Rajivan
School of Informatics & Computing, Indiana
University
Bloomington, IN, USA
prajivan@indiana.edu

Jean Camp
School of Informatics & Computing, Indiana
University
Bloomington, IN, USA
ljcamp@indiana.edu

ABSTRACT

Transmission of personally identifiable information from smartphone apps has become ubiquitous as smartphones themselves. Privacy controls currently provided in the form of permissions warnings falls insufficient especially for communicating risk during app installation. Priming for privacy and showing easy to understand risk cues could help people make low risk app choices. Towards this, we conducted a user experiment with 480 participants who made a series of app choices with/without privacy priming and with/without risk communicating cues. Overall, presenting risk communicating cues along with app benefits has a significant effect on user app choices in terms of risk-benefit tradeoff. We found that priming for privacy would lead to increased concern while choosing apps but may not have a strong effect on final app choices when combined with certain types of risk cue framing. We conclude with recommendations for improving risk communication in Android.

Author Keywords

Risk Communication; Privacy; Privacy Priming; Usable security; Human Factors; Smartphones; Android; Simulation

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation(e.g., HCI): User Interfaces; D.4.6 Security and Protection: Access Controls

INTRODUCTION

Smartphones have now become a source of continuous personal data for businesses, advertising services and even malicious third parties. Much of this data stream originates from apps and data-driven, personalized services running on smartphones and includes personally identifiable information such as location information, search queries, personal messages, personal media (e.g photos and videos), health information (health tracking information) and financial information. While governmental and corporate location tracking and data compilations (such as with metadata) have led to widespread distress, many users are unaware of the amount of personal data sent to unknown third parties from their own smartphones. High profile events, such as the indictment of the CEO of Stealthgenie [31], sometimes bring these issues to the fore. However, such apps' based privacy risks still remains largely vague to users making app installation decisions on a day-to-day basis.

Surveys have long indicated that individuals value their online privacy [36] but research about users' beliefs with respect to privacy implications of their own behaviors shows a disconnect between belief and reality [25, 24]. Thus, not surprisingly, research focused only on behavior shows that people do not act in a privacy-preserving manner which is inconsistent with their expressed preferences. One of the reasons for such a differential user behavior is information asymmetry and users' uncertainty about the data collected by software services in the background [2] making. People have also voiced their discomfort towards existing data collection activities on smartphones. In one research, when users were given explanations about permissions warnings and how apps installed on the phones were using their personal information, they expressed severe concerns and discomfort [38].

Desirable personalization through data compilations is common in smartphone applications. However, usage of personal information in direct opposition to users' privacy preferences is also common in smartphone world. Access to personal information by apps has been found in both iOS [18, 4] and Android [21, 10, 20, 44, 5]. Disclosure of information for obtaining personalized services can therefore be characterized as privacy bargain ideally made through risk/benefit tradeoff [1, 30, 40]. In some cases, such a bargain is necessary, benign, and even desirable but in other cases it could lead to invasion, discrimination, and even exclusion [2]. Making an informed, risk aware app installation decision based on risk/benefit trade-off analysis poses as a difficult task for general users because risk and benefits of Android apps are often intangible and asymmetric [26].

Currently, to address privacy concerns, smartphone OS providers are leveraging privacy frameworks and standards to ensure that user's information is aggregated and stored in a privacy preserving manner [41]. (In this paper ”OS providers” refers to Apple for iOS, Google for Android, Microsoft or WINDOWS, and Amazon for FIRE). Furthermore, OS providers are taking efforts to encourage their third party app developers to inform and respect consumer preferences. For example, Apple has recently started pushing their app developers to exclusively use HTTPS connections for new iPhone applications. This is in response to concerns about privacy of data in transit. Even though OS providers are taking steps to improve user privacy which includes evaluation of third party apps to detect malware and obvious privacy concerns, it is however not clear (and admittedly difficult to ascertain) whether app developers are using similar kinds of privacy standards in developing their apps and therefore present a major concern for smartphone user privacy. Hence when risks cannot be effectively detected and completely mitigated, it is imperative to effectively communicate the risk to users to help them make informed app choices.

Different OS providers use different approaches to inform customers about privacy risks. Google in their Android OS used to present a complete list of permission warnings after app selection and before installation. The list is entirely based on the information in a manifest provided by the developer and are not inspected for validity. However with Android 6.0, Google has adopted Apple iOS like permissions model wherein permission warning prompts are presented at the first access to a certain resource by an app. Such run time warnings are known to by affected by user habituation effects to ignore and click through warnings [43, 19]. Neither permissions-based model has proven to be effective in communicating risk to the users [4].

For usable and secure mobile systems, it is imperative that risk is communicated only when user intervention is absolutely necessary (when risk is unavoidable or when user preference/decision is required to proceed). When necessary, risk must be communicated early on in the decision process along with other relevant decision variables and presented to the user in a manner that is easily comprehensible even for security naive users. Following these simple, user-centric principles would empower all users to make quick and low risk app choices with little to none uncertainty. More risk warnings during and after installation will only lead to habituation and consequently disregard from users.

In this paper, we present our work on influence of privacy priming and risk communicating cues on user risk behavior during app installation in Android. We specifically investigated the interacting effects of different risk framing and privacy priming on user app choices measured in terms of risk-benefit trade-off. We also present our work that explored the privacy paradox in Android app choices i.e., disconnect between stated privacy preferences and actual privacy behavior during app installation.

RISK COMMUNICATION IN SMARTPHONES

Ideally systems must be engineered without risks to humans but that is often not feasible. When elimination of risk is not feasible, effective warnings about risks must be shown to elicit the desired behavioral response from humans [16]. Likewise, eliminating risk to privacy for smartphone users is not feasible but effective warnings can be provided to help users make risk averse decisions while using their smartphones. Several frameworks has been proposed to guide warning development of which the communication human information processing framework (CHI-P) [16] presents a high level 5 stage framework that structures how warning information would be processed by a human receiver and how the effectiveness of processing is vital to stimulate a compliance behavior in human receivers. The stages in the CHI-P framework include Attention, Comprehension, Attitudes and Beliefs, Motivations and response behavior. Failure at any of these stages can become a bottleneck that inhibits the desired behavioral response.

It is necessary that Android users pay attention to permission warnings, read and comprehend them to consequently make risk-aware app installation decision but past research has shown that people usually ignore or pay little attention to the permissions warnings in Android [23]. It was found that only 17% of participants self-reported to have paid attention to the permissions and 42% of participants were unaware of the existence of permissions. One of the primary reasons for such a user behavior could stem from users' habituation to warning dialogs wherein users have habituated to ignore digital warnings in general [39, 43, 19, 12]. Similarly, due to habituation, past research has also found that users click through most of the smartphone permission warnings and therefore gain no awareness about the resources being used by the apps [14, 11]. Such a lack of awareness is not constrained only to Android based phones but also in other smartphone operating system platforms such as Apple iOS [35].

In addition to users' inattention towards permissions warnings, past research have also found that Android users do not fully comprehend the permissions presented to them during the application installation process [23]. They found that only a few participants carefully read the permissions requested and made decisions about whether to install an app or not based on the permissions requested by the application. One reason for such a seemingly risk seeking behavior is that majority of smartphone users are security naive and therefore are not qualified to make an informed decision considering all the security risks. Hence they need simpler, easy to comprehend, visual cues to assist me in making risk averse app choices.

Variance in peoples' literacy in terms of basic reading literacy, English literacy and computer technology literacy could serve as significant factors that amplifies privacy risk comprehension problem in smartphones. App permissions are commonly requested in English with too much jargon. Internationalization and localization features offered by the makers do not extend beyond a handful of languages and is often not configured appropriately by the app developers. It is invalid to assume that all smartphone users across would possess an above average level of literacy along with computer/mobile technology knowledge to comprehend the permissions information presented to them to predict the implications of agreeing to the requested permissions. Therefore it is imperative to communicate smartphone privacy risks using simple modes that do not demand high language and computer literacy from users.

Strong motivation and attitude to make all app choices consistently based on privacy risk may not be feasible for all smartphone users. App choice has become an everyday activity and so people are not mentally motivated to make risk aware app choices even when they warned for risks. Therefore, it is necessary to stimulate people's memory for privacy risks to mentally prepare them to make risk aware app choices. Priming people for privacy risks is one way to approach this problem. We hypothesized that priming people for privacy risk will have a significant effect on user app choices.

Framing, Heuristics and Biases

People are minimalist in using cognitive resources. Hence we use heuristics to simplify our everyday decision making process but the downside is that our decisions could be rife with cognitive biases. Prospect theory [42] is a model of user choices under uncertainty which demonstrates preconditions that lead to either risk averse or risk seeking human behavior. Kahneman and Tversky empirically showed that there is a asymmetry in the way humans make choices wherein humans underweight outcomes that are probable in comparison to definite outcomes. People choose to make risk averse decisions when presented with choices that contain outcomes with certain probability of gain. For example, it was found that when offered a choice between gaining $1000 with certainty and gaining $2500 with 50% chance, people more often chose $1000 (risk averse decision). In contrary, people choose to make risk seeking decisions when presented with choices that contain outcomes with certain probability of loss. For example, it was found that when offered a choice between loosing $1000 with certainty and loosing $2500 with 50% chance, people more often chose to loose $2500 (risk averse decision). Likewise, on smartphones, people make app choices predominantly based on definite benefits offered by the app while giving less consideration towards privacy risks from the usage of risky apps.

Based on how the choices are framed [42], peoples' decisions vary wherein certain framing could lead to making risk averse decisions while others could lead to risk seeking behaviors. In general, such a phenomenon is called a "framing effect" when different but equivalent representations of the same choice quandary lead to different decision outcomes. Similarly, we can hypothesize that different risk framing would lead to different user risk behavior while choosing apps.

RELATED WORK

Past work on risk communication in Android has predominantly focused on improving the permissions interface through simplified text [9], explanations with more text [29], explanation of warnings with examples [28] and some through visual cues [9, 15, 13] that represents the threat level of the permissions requested.

Past research has also looked at presenting permissions in the App description page in Android instead of presenting it after the users choose to install an application [29]. This research showed that displaying permissions early on before the users make up their mind about purchasing an app could improve application choices in terms of risk. However in this research, the subjects were asked to imagine that they were choosing the apps for a friend and from a methodology perspective, it is potentially incorrect to present a scenario where the participants are asked to choose apps for a loved one or a friend because it then measures risk recommendation and not risk perception and that people have been found to be more impartial and risk averse while recommending a risky situation to others but at the same time can be emotional, biased and risk seeking when it is for themselves [27].

There has been a limited past exploration of effects of risk framing on App choices. In one research [15], subjects were asked to compare positive framing of risk communication against negative framing and found that both positive and negative framing did not vary significantly. However in this study, the subjects were making a comparison between two scales without any context. In another experiment reported by the same authors, they found that visual cues could have an effect on how users make risk based app decisions. However the effect was measured by presenting subjects with the same app repeatedly and by asking to make a comparison between the two scales (negative and positive). Similarly, in another research [13] subjects were asked to make choices either based on positive framing or negative framing of risk and found that subjects made better choices with positive framing in comparison to negative framing. However, the experiment did not have a control condition to measure and compare the effectiveness of presence of risk cues against no visual risk cues.

Past research on risk communication in Android has mainly focussed on improving permissions page and have not given much importance to the effects risk communication early on in the decision making process. Some of the past work on risk communication was also found to have been conducted using experimental scenarios that could have potentially confounded the results. There is limited to none past work on the influence of motivation in Android app choices. No research has been done to explore the extent/strength of the effect of visual risk cues on app decisions.

To fill this critical gap, we conducted an Internet-based user experiment. In our experiment, we studied the influence of privacy priming on app choices, measured and compared the effect of visual risk cues to no cues and through repeated trials, we measured the extent of risk framing effects on app choices. We developed a Android v5.0 (current version of Android when the experiment was conducted) App store simulation system to measure user app choices and the associated experiment variables. Participants used the simulation system to make a series of Android application installation choices after which they responded to questionnaires on Android usage, Android permission comprehension, security knowledge and demographics.

METHOD

Android Play-store Simulation

A fully interactive Android play store simulation was developed for the experiment. It was a simulation of play store as in Android v5.0 (a.k.a lollipop) in which the list of app permissions is displayed just before app installation. At the time of experiment, Google's new permission model in Android v6.0 (a.k.a Marshmallow) was not unveiled and therefore was not explored. This system (see Figure 1) was a simulation of play store interfaces and navigation capabilities used in exploring and installing apps on Android based phones. Goal was to develop a realistic simulation system that requires realistic user interactions and task-flows which in turn would require participants to exercise some of the same cognitive process involved while choosing apps in the real world [17].

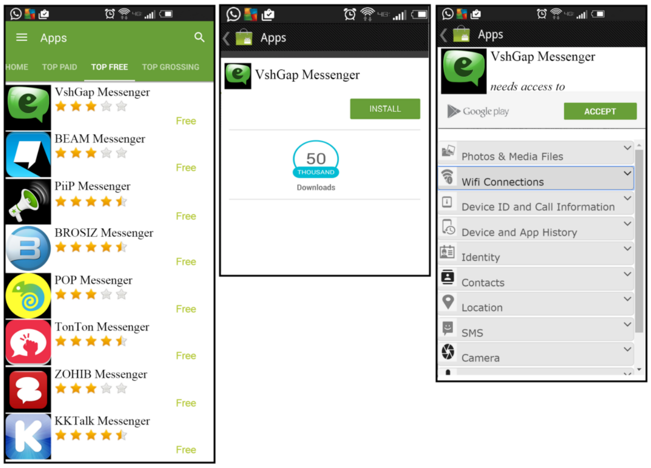

Figure 1: Screenshots of the three simulated Android app-store interfaces

The simulation system was designed using actual images of Google play store and images used by real apps on play store. All text, font, and status messages (e.g., "installed" status for apps that was installed) were closely matched to the layout of play store interface. The simulation included three main screens involved in app installation decision process: page with the list of apps, individual app description page and app permissions page. Similar to real world play store, the first page on simulation showed a list of apps belonging to a certain category. On selecting an app, the second page with corresponding app details and download count was displayed. The download count information using the same icons used by Google was retained in the app description page in addition to app name and image. Finally, when the user chose to install an app, the corresponding permissions page was displayed. At this point, the participant had two choices, to either install the app or go back to the list of apps. On choosing to install, the participant was taken back to the screen containing list of apps but with a status message "installed" appearing next to app that was chosen as appearing in the actual play store interface. Participants, during the experiment instructions phase, were informed that they were not actually installing the app and that this was a simulation. Effort was put to match the permissions page to closely replicate the app permissions page displayed on Android v5.0 version. It replicated the summary, collapse and expand appearance used on play store in Android v5.0. The icons and text used to represent the different permissions was also replicated in the simulation. The interactions were also retained wherein by default all permissions in the page was in the collapsed state and the participants would explore them by expanding each one similarly to the play store interface.

Furthermore, the permissions for a category of apps was based on actual permissions requested by apps in that category. The action of expanding a particular permission was recorded to measure the types of permissions explored by users in each category of apps and experiment condition. The system was fully interactive such that the participants had the ability to go back and forth between pages as they would do on Android phones. They were also allowed to change their choices and uninstall the previously installed apps. The participant could uninstall and install other apps in a particular category until he/she moves on to make choices in the next category of apps.

For experimental control and to reduce confounds arising from differences in screen size and layout, participants were instructed to perform all the experimental tasks using an Internet browser on a desktop/laptop machine with the simulation interface centered and scaled to the size of a standard large smartphone screen dimensions. Furthermore, only native Android users or users with Android OS experience was encouraged to participate in the experiment to reduce confounds that might stem from lack of familiarity with Android OS. This was achieved by using a questionnaire based screening process at the start of the experiment.

Apps

Apps belonging to eight categories of apps were used in the simulation and they were: "Password Manager", "Ebook Reader", "File Manager", "Messenger", "Puzzle Games", "Fitness", "Dating" and "Photo Editor". For each category, eight number of apps were presented to the participant. Apps used in the simulation were actual apps from Google play store. Apps that were ranked 75 and above on the Google play store at the time of experiment was chosen for simulation to reduce biases in choices due to app popularity and recognition.

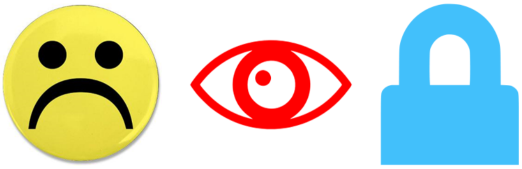

Figure 2: Screenshots of the three risk cues used in the experiment

Framing of cues for communicating privacy risks

We hypothesize that, cues that are attention grabbing, easy to comprehend, literacy neutral and conforming to peoples' mental models are necessary to effectively communicate smartphone app privacy risk to users. As shown in Figure 2, we explored three types of framing of cues (Emoticons, Eyes and Padlocks) that intuitively satisfy the stated conditions.

Emoticons are popularly used to communicate emotions on text-based communication media. Social cues such as eyes was explored in this study because it has been found in the past to elicit user reactions to risk warning [9]. Finally, cues that represent physical security (padlocks) which has been shown to conform to peoples' mental models about online security was also explored in the study. The cues were represented either in risk scale (negative framing) or privacy scale (positive framing). In risk scale framing, higher values on the scale represents higher risk and lower values represents lower risk. On the other hand, in privacy scale framing, higher the value on scale represents higher privacy and lower the value on scale represents lower privacy.

Privacy Priming

Privacy priming in this experimental context was done using a short and modified version of the IUIPC (Internet Users' Information Privacy Concerns) questionnaire [33]. Text or video based descriptions to prime for privacy concerns was not used because memory recall are better through test like instruments [37] such as the IUIPC scale. Since the goal was to prime the users for privacy, only eight privacy concerns questions from the IUIPC instrument was asked. The eight questions asked are as follows:

- All things considered, the Internet would cause serious privacy problems
- Compared to others, I am more sensitive about the way online companies handle my personal information
- To me, it is the most important thing to keep my privacy intact from online companies.
- It is very important to me that I am aware and knowledgeable about how my personal information will be used
- Compared with other subjects on my mind, personal privacy is very important.
- I am concerned about threats to my personal privacy today
- It usually bothers me when online companies ask me for personal information.
- When online companies ask me for personal information, I sometimes think twice before providing it

Experiment Variables

For each app choice the participant made, three dependent variables were measured in the simulation part of the experiment. They were app rating, download count and the privacy/risk score of apps. App rating (aggregated user rating of apps) and download count (number of times an app was downloaded from the play store) is representative of the standard app metrics predominantly used by people to currently select apps. Privacy/risk score is the proposed additional metric for apps. Risk score in this experimental context communicates the level of risk to personal information and privacy score communicates the level of privacy offered by the app. Other variables measured in the experiment include measures of user attention towards Android permissions and measures of Android permission comprehension. The results on permission comprehension are not presented here as it is beyond the scope of this paper.

Experiment Design

The experiment was a 4X2 between subjects experiment design. The framing of visual cue to communicate privacy/risk score was one of the independent variable and it had four levels. The four levels of privacy/risk score cue includes no cue, frown face, human eyes and padlock. Privacy priming was the other independent variable and it had two level: Privacy primed or not primed. No visual cue to communicate privacy/risk was the control condition and is representative of the current system. Frown face communicated risk and was emoticon based cue that is widely understood by smartphone and Internet users. Eyes also communicated risk and was a cue for eliciting response through social cognition [3]. Finally, padlock communicated privacy and was a security mental models based cue [6].

Experiment Procedure

480 participants from Amazon mechanical turk participated in the study of which 233 were females and 247 were males with an average 31 years of age. Participants at the start of experiment (after accepting HIT on MTurk) responded to questions about their age and familiarity to different smartphone operating systems (such as iOS, Android and WINDOWS). Participants who self-reported their age to be above 18 and who self-reported to be familiar with the Android OS were only allowed to participate in the experiment. All participants in the study were above 18 years of age. Participants were then presented with information about the study and was informed that taking part in this study was voluntary. Participants, on providing consent to participate, were randomly assigned to one of the 8 experimental conditions such that there were 60 participants in each condition at the end of the study. On completion all participants were paid $2.50 for their participation in the experiment. The study procedures were approved by our university's Institutional Review Board (IRB) for human subjects' research.

Based on the experimental condition assigned, participants were either presented with the privacy priming questionnaire at the start of the experiment or at the end of the experiment. In the priming condition, these questions were presented before participants made app choices on the simulation and in non-priming condition, these questions were asked after participants made app choices. As you would recall, privacy priming questionnaire was used to prime them about the importance of online privacy and to stimulate their online privacy awareness. Privacy priming was done through the short IUIPC privacy questionnaire [33]. Completion of this questionnaire took the participants to the next step in the experiment which involves using the online simulation environment to make application choices. Prior to using the simulation environment, participants were provided specific but simple set of instructions on using the simulation environment. Since only native Android users were encouraged to participate, the learning curve to get accustomed with the system would be short.

During the simulation, the participants were presented with eight categories of apps with eight apps in each category displayed in a randomized order. Each category of apps were

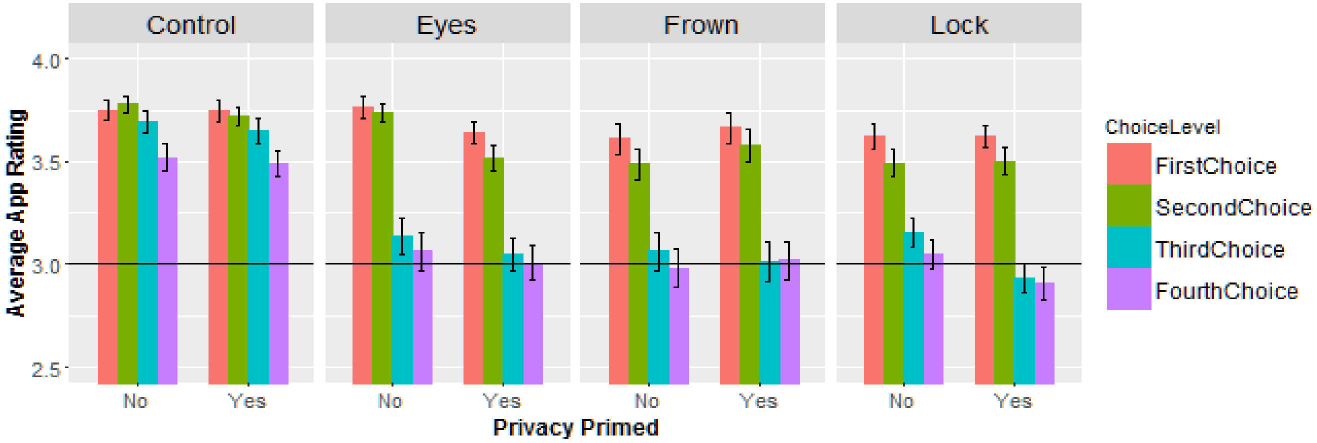

Figure 3: Mean values of app rating of choices across the eight experimental conditions. Horizontal black lines are for reference.

presented to the participant one after another. Two levels for each of the three experiment variables: app rating, privacy/risk score and download count, were used as attributes for apps in the simulation environment, making the possible number of variable combinations to be 8 (2^3). The 8 combinations of the variable set app rating, privacy/risk score, download count) were randomly assigned to the 8 apps in each category such that each app in a particular category had exactly one of the 8 possible variable set combinations. App rating and risk score ranged between 1 and 5 with two levels: 2 represents low-medium app rating/low-medium risk score and 4 represents high-medium user rating/high-medium risk score. Similarly two levels of download count was used: fifty thousand and hundred thousand downloads. Very low or very high app rating or risk score values such as 1 and 5 were not used. Similarly very low or high download values such as in hundreds or in millions were avoided. This was done to ensure the attributes associated with apps were realistic and also to recreate the complexity of making risk benefit trade-off decisions with real world apps. Participants were instructed to choose 4 of the 8 available apps to install in each category. After the selection of 4 apps in a particular category, the participant had the option to move to the next category containing next set of apps. There was no rigorous time constraints placed on participants to make the four app choices in each category. Participants were given upto an hour to complete all the experimental tasks.

Participants were specifically instructed to choose 4 apps in each category because only 4 of the 8 combination of variables has atleast two variables with desirable values (e.g., high app rating, low risk or high download count) and the remaining 4 apps would have only one or none of the variables with a desirable value. Hence, having participants select 4 apps as opposed to choosing one or two apps per category would help in measuring the extent of effect of cue and/or priming on app choices. With each app choice, the availability of good app choices reduces leading to choice complexity especially while choosing the fourth app. If the effect of cue and/or priming is strong, it can be hypothesized that the participants would choose all four apps with desirable risk and benefits values. After participant made all the required application choices for all 8 categories, they were presented with a series of questionnaire to measure their comprehension on the textual description of the permissions as presented in Android v5.0 and then a set of questionnaire to capture their demographics such as age, education and income.

Data Collected

Several implicit and explicit measures were collected from the experiment which includes app choices made in each app category, permissions viewed on each app, amount of time spent on choosing apps in each category, responses to Android permission comprehension questionnaire, self-reported android permission behavior data and responses to online privacy questionnaire (used to prime for privacy). As was stated before, experiment participants chose 4 apps from each of the 8 categories of apps presented to them (total 32 app choices). The apps selected by participants were recorded in terms of the privacy score, app rating and download count values associated with the chosen apps. Results from Android permission comprehension questionnaire, privacy questionnaire and android behavior questionnaire is not presented in the paper as it is beyond the scope of the paper.

RESULTS

We received responses from almost equal percentage of males (51.35%) and females (48.65%) between the ages of 18 to 75. The average age of participants was 31 years. Data from 42 participants were excluded from analysis because they either chose only the top 4 appearing apps for more than 4 categories and/or completed the entire study in a very short duration (under 5 minutes). These metrics were used as indicators for participant's lack of cognitive effort put towards making app choices. Analysis was conducted on data from the remaining 438 participants.

The three app variables (app rating, privacy/risk score and download count) were originally recorded on different scales in the simulation. Therefore, values for all three variables measured from all eight experiment conditions were normalized to be in the same range of 2 to 4 with 2 representing

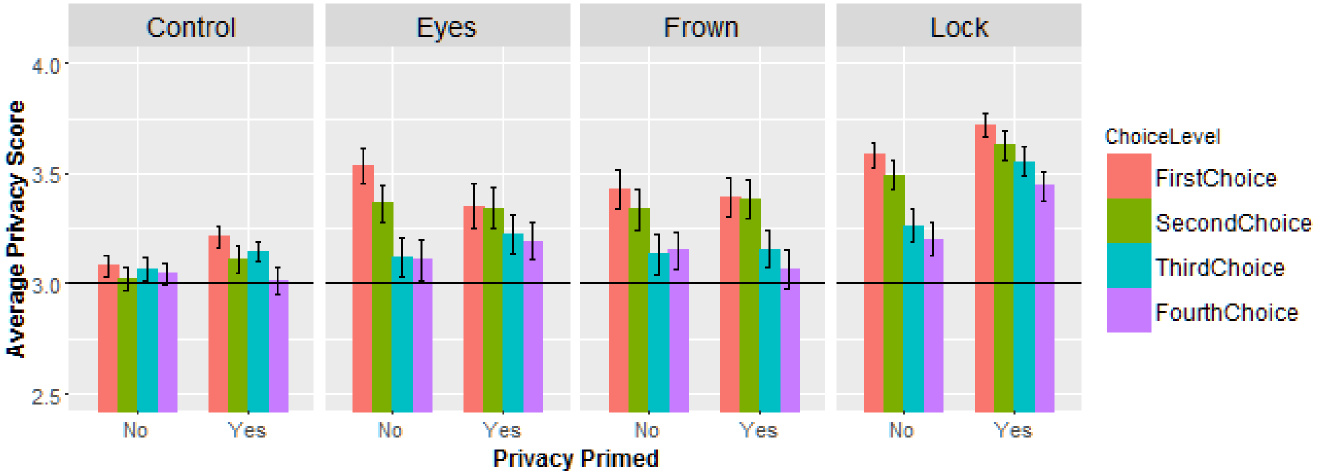

Figure 4: Mean values of privacy score of choices across eight experimental conditions. Horizontal black lines are for reference.

app choices with low-medium app rating/low-medium privacy score/low-medium download count and 4 representing app choices with high-medium app rating/high-medium privacy score/high-medium download count.

App choices were measured and compared based on the three app variables (app rating, privacy/risk score and download count) at each of the 4 choice levels (First, Second, Third and Fourth choices). Such a choice level analysis was conducted to measure the strength of the effect of risk communication cues and privacy priming on choices. Towards that, for each participant and at each choice level, the privacy score, app rating and download-count values of chosen apps were averaged across the 8 categories of apps. Since values of variables were either 2 or 4, on averaging, the mean values close to 2 and 2.5 (floor) are app choices with low app rating or less privacy offered (or high risk) or less downloaded whereas mean values close to 3.5 and 4 indicates app choices with high app rating or high privacy offered or more downloaded and values around 3 indicates inconclusive/indeterminate choices.

Figure 3 is a graph of mean values of app rating for all four app choices in all eight experimental conditions. As it can be seen, app rating of all four app choices in the control condition (when primed or not primed for privacy) was between 3.5 and 4 which indicates a strong influence of app rating in all four choices in the control conditions (when there are no cues to communicate risk). In conditions containing risk cues (both risk and privacy framed cues), app rating of the first two app choices was between 3.5 and 4 indicating the influence of app rating whereas the mean app rating of third and fourth app choices in these conditions were close to 3 indicating inconclusive effect of app rating on third and fourth app choices. Mean values of download count for all four app choices in all eight experimental conditions was analyzed. Mean download count of chosen apps was found to be close to 3.0 across all eight conditions indicating inconclusive effect of download count on all app choices. Finally, Figure 4 is a graph of mean values of privacy score for all four app choices in all eight experimental conditions. As it can be seen, mean privacy score for all four choices in the control condition (both primed and not primed for privacy) is close to 3. Although, the privacy score of choices in primed control condition is marginally higher than privacy score of choices in non primed control condition. In comparison, mean privacy score of app choices in the three conditions with risk communicating cues has a lot more variability between conditions. Specifically, in the condition which used padlock for risk communication and also primed participants for privacy, the privacy score of all four app choices was close to 3.5 indicating a strong influence of privacy score on app choices. In comparison, in other conditions with risk cues, we don't observe such a strong influence of privacy score on all four app choices.

The three variables (app rating, privacy score and download count) that describe the app choices were found to violate normality assumptions (Shapiro-Wilks W<0.96 p<0.01). Hence, the non-parametric alternative to ANOVA called Kruskal-Wallis test was employed to measure and compare the effect of risk communicating cues and privacy priming on app choices. Since, Kruskal-Wallis test cannot be applied to a factorial design, the 4X2 between subject factorial design was transformed to perform one-way non-parametric ANOVA test between 8 experimental conditions: Control-Primed, Control-NonPrimed, Frown-Primed, Frown-NonPrimed, Eyes-Primed, Eyes-NonPrimed, Lock-Primed and Lock-NonPrimed. Kruskal-Wallis test **on app rating** revealed that there was no significant effect of risk communicating cues and privacy priming on the first app choice across the eight experimental conditions (χ^2 = 11.3, df=7, p = 0.12). A significant effect of risk communicating cues and privacy priming was observed on app rating on the second app choice (χ^2 = 20.9, df=7, p <0.01), third app choice (χ^2 = 83.2, df=7, p<0.01) and fourth app choice (χ^2 = 55.4, df=7, p <0.01).

Kruskal-Wallis test **on privacy score** revealed that there was a significant effect of risk communicating cue and privacy priming **on privacy score** on the first app choice (χ^2 = 63.9, df=7, p < 0.01) second app choice (χ^2 = 49.7, df=7, p <0.01), third app choice (χ^2 = 28.7, df=7, p <0.01) and also fourth

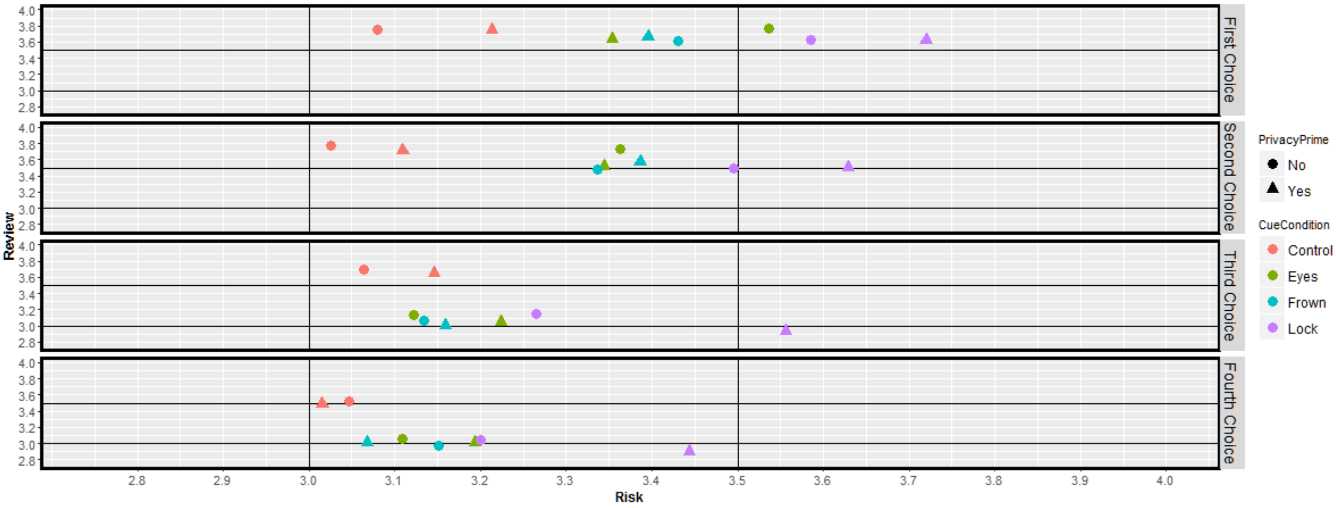

Figure 5: Risk Benefit graph of choices Across the eight experimental conditions.

app choice (χ^2 = 21.9, df=7, p <0.01). Post-hoc Tukey tests revealed that there was a significant difference in privacy score in first and second choice between control conditions (both primed and non-primed) and all 6 remaining conditions with cues. In the first and second choice, the privacy score of apps in the control condition is significantly smaller in comparison to other experimental conditions. Post-hoc Tukey tests also revealed a significant difference in privacy score in the first choice between primed and non-primed control conditions with privacy score in primed condition being significantly higher than the non-primed condition. Finally, posthoc tests also revealed a significant difference in third and fourth choices between the Lock-Primed condition against all other 7 experimental conditions with the privacy score in lock-primed condition being significantly higher than the other conditions. There was no significant difference between any other two conditions in third and fourth choice in terms of privacy score.

Kruskal-Wallis test **on download count** variable revealed that there was no significant effect of cue and priming **on download count** on the first (χ^2 = 1.8, df=7, p= 0.96) and fourth app choices (χ^2 = 4.7, df=7, p=0.68). Although, a significant effect of cue and priming on download count on the second (χ^2 = 15, df=7, p=0.03) and third app choices (χ^2 = 18, df=7, p=0.01) was detected. Median download count values in all conditions for all four app choices was found to be close to 3 (inconclusive) which rendered the download count variable unusable for further analysis. Based on results from non-parametric tests, it can be inferred that participants were making a risk-benefit analysis predominantly using the app rating and privacy score and that download count of the app did not play a significant factor in their decisions.

Figure 5 is a graph of app choices in terms of risk and benefits. In Figure 5, the x-axis is risk (in terms of average privacy score of app choice) and y-axis is the benefits (in terms of average app rating of app choice). Each panel in Figure 5 is a choice level. Vertical and horizontal black lines are for reference drawn at units 3 and 3.5. As before, average value of 3 indicates inconclusiveness (values are inconclusive/indeterminate on whether they are risk or benefits based or both). As it can be seen, participants in the control condition (points and triangles in orange) made their choices predominantly based on benefits (app rating) because the points lie around the X-Y coordinates (3.0,3.5) in all four choice levels. Although, risk appears to have played a faint role in first choice in participants in control-primed condition (control condition who were primed for privacy ,denoted by orange triangle). First two app choices in all experimental conditions containing some form of risk communicating cues appears to be based on both risk and benefits because the points lie close to the X-Y coordinates (3.5,3.5). However, in third and fourth choices, only participants in the lock-primed condition (condition using padlock and were primed for privacy, denoted by purple triangle) appears to be making app choices predominantly based on Risk because points lie around X-Y coordinates (3.5,3.0). The third and fourth choice in all other conditions containing some form of risk communicating cues appears to be inconclusive because the points lie close to the X-Y coordinates (3.0,3.0).

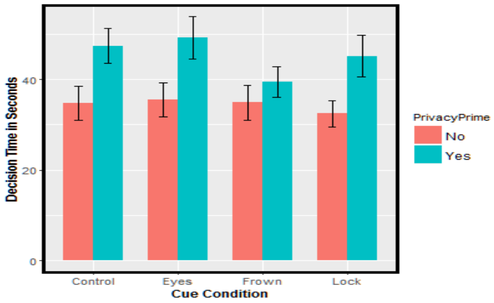

Figure 6: Average time taken to make app choices across the eight experimental conditions

Time taken to make app choices was measured. Preliminary analysis showed that participants spent most amount of time in making the first choice which in turn indicated that participants first made all four app choices inside their head and later made actual app installation one by one. Therefore, time taken to make the first choice was assumed as indicative of time taken to decide all four app choices. Average time taken to make the first app choice was compared across the eight experimental conditions as show in Figure 6. As it can be seen, irrespective of the type of risk communication cue, participants when primed for privacy spent significantly more amount of time in making app choices (20 seconds more in average).

Both self-reported and actual users' attention towards Android permission was measured. After making all the app choices on the simulation system, participants were asked how often they reviewed or read the permissions presented to them while they install applications from Android Play Store. As shown in Figure 7, 3.96% of the respondents said they "Never" read the permissions, 43.96% of the respondents said they read the permissions "Sometimes", 36.25% of the respondents said they read the permissions "Almost All the time" and 16.46% of the respondents said they read the permissions "Everytime". As an implicit measure of users attention towards permission, we recorded the total number of times each participant clicked and opened a permission on the permissions page across all 8 app categories of apps. Figure 8 shows the frequency distribution of such a count measure. As it can be observed, 77.92% of participants did not click open even one permission across the 8 categories of apps. There was no significant difference detected in terms of user attention to permission between the experiment conditions. This is in direct contrast to the self reported permission behavior described earlier.

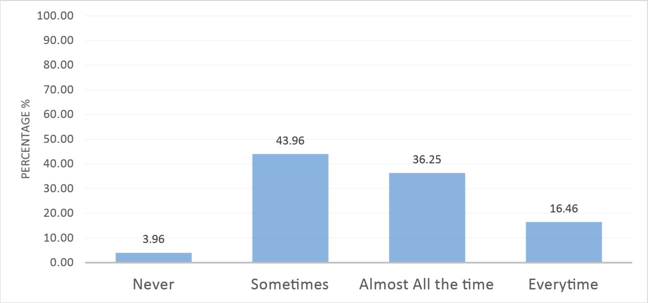

Figure 7: Self-Reported frequency on permission warning review

DISCUSSION

We aimed at measuring the influence of risk communicating cues and privacy priming on Android app installation decisions. Towards that objective, we measured app choices in terms of three app features: user rating of App, download count of app and privacy score of app. Furthermore, to explore privacy paradox in Android app choices, we measured user attention towards Android permissions at install time in terms of number of permissions actually viewed by participant in the simulation environment and compared it against the expressed Android permission behavior.

Consistent with past research [42], we found that app choices in Android can be biased depending on the information available to people early on in the app installation decision process. Participants in the control condition (with no risk communicating cues) made app choices predominantly based on app rating while the influence of risk information (Install time Permissions request) on their app choices was found to be minimal to none. It can be inferred that it is easier for people to rely on app rating to make app choices when it is difficult to assess app risk through privacy permission disclosures as shown in Andorid. In contrast, when participants were presented with risk communicating cues alongside app ratings, participants were found to make better risk averse app choices by incorporating both risk and benefits information. It was found that download count information did not influence app decision making process. Part of the reason could that it was displayed only when an app was opened for further description and was not presented alongside other variables such as app rating and privacy score. The other reason could be that the download count values used for the experiment (50 thousand downloads versus 100 thousand downloads) may have been not discriminating enough to have an influence on app choices.

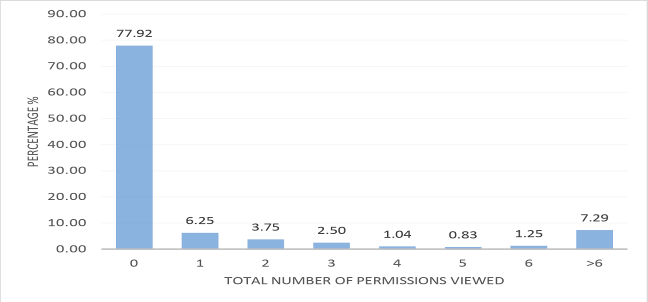

Figure 8: Actual number of permission warnings viewed by participants

Framing privacy risk information differently was also found to have a significant effect on app choices wherein framing risk in terms of privacy offered, lead to better app choices than framing in terms of risk to one's privacy. When participants made choices using cues with risk framing (frown face and eyes), their first two app choices (when there are more app choices to choose from) were based on both risk and benefits information. However, after the first two choices (as the number of available options reduced), the effect of risk communication on app choices was found to deteriorate and started to look stochastic (Inconclusive) wherein it was not clear whether decisions were based on app rating or risk score or both. When participants were presented with cues communicating privacy offered by the apps (using padlocks), decisions on the first two app choices were based on both risk and benefits, similar to results in other conditions with cues. In contrast, participants making choices using cues with privacy framing, made the third and fourth choice predominantly based on risk indicating the strong influence of privacy framed risk cues on app choices over other variables such as app rating. One explanation for such a strong effect of privacy framed cues is that the privacy score was displayed on the same scale as app rating and therefore did not require mental rotations from people while incorporating risk and benefit information in making decisions. It could also be hypothesized that the privacy score was communicated using padlocks which has been found to be representative of user mental models of security [6].

Overall, priming for privacy was found to have less influence on app choices in comparison to the effect from risk communicating cues. Participants in the control condition (with no risk communication), when primed for privacy, was observed to make their first choice incorporating both risk and benefits information. However, the effect of priming was found to deteriorate beyond the first choice. After the first choice, participants in the control condition fell back to making choices predominantly based on app rating. Effect of priming was also observed with participants making app decisions using padlock cues for risk information (privacy framing). It was observed that privacy priming in this condition alone enabled participants to make risk based app choices even in the third and fourth app choice levels whereas participants making decisions with same type of cue but who weren't primed for privacy, were making inconclusive app choices. Finally, time taken to make app choices indicated a strong effect of privacy priming. When participants were primed for privacy, they took significantly more amount of time in making app decisions (20 seconds more in average). It could be theorized that priming for privacy could be leading to increased concern in participants but without appropriate risk communicating cues, people are unable to understand the risk implications of apps, failing to make consistent risk aware app choices. However, when paired with appropriately framed risk communicating cues (such as padlocks), priming for privacy can augment user risk behavior and therefore leading to more consistent risk averse app choices. These results show the important role of motivation and attitude towards privacy in smartphone app choices.

Almost 97% of participants self-reported that they read/reviewed Android permissions when requested. However, in reality, we found that participants paid very low attention to permission warnings. 71% of participants did not view (by expanding it on the interface) any permission warnings across the 8 categories of app they viewed (64 apps in total). Either they completely ignored the warnings, or simply glanced over the warnings page. The other possible explanation to this result is that they knew the implications of permissions requested and therefore did not see the need to open and view them. These results strongly complements past research findings [23] about users' inattention towards Android permissions warning page. Such inattention to permission warnings could in part be because of the sheer amount of time and cognitive resources needed to review and make decisions based on permission warnings. Moreover attention is a limited cognitive resource that is applied only to environmental cues that are relevant. Furthermore, these results provide evidence for privacy paradox in android app choices, because there is a distinct difference between people's expressed Android permissions behavior and their actual privacy behavior.

From this experiment, we observed that end-users currently don't understand permission warnings and are not paying enough attention to permission warning. Hence they are often making app choices based on benefits ignoring the risks. Communicating privacy risk through easy to understand cues could help people make informed app choices. Results from this work indicates that cues and scales for communicating privacy risk must be carefully chosen to help users make informed app choices because presenting risk/privacy score in association with other metrics such as app rating can lead to non-intuitive app choices and that different cues and scales can have different degrees of effect on app choices. Results from this work are directly applicable to privacy engineers, designers and researchers designing ways to effectively communicate smartphone privacy risks. Communicating privacy risks using literacy neutral, easy to understand cues could help all users to make informed, risk averse choices which would increase the cost of over-declaring permissions for app developers [7, 8] which in turn would discourage developers from over-declaring permissions and would therefore encourage them to access just the required amount of user information.

Privacy priming is essential as users would need to be reminded and motivated to be risk aware. Experiment participants were primed for privacy using the IUIPC [33] scale. Hence, future work would need to identify more simple and practical ways to prime users for privacy. The risk metric used in this experiment was simply based on the number of permissions requested by the app which is not an accurate risk metric but served the purposes of this experiment. There is a significant amount of past work in taint analysis [21, 10, 20, 44, 5] which can analyze Android apps (using Android Application Package) with reasonable accuracy to detect taints in the form of information leaks. Outputs from such code analysis tools can be leveraged to develop reliable risk metrics for apps and communicated using results from this work. Presenting risk communicating cues along with app ratings could potentially lead to incorrect sense of safety if the methods used for measuring risk is incorrect or simply an approximation. Hence measure of app risks must deployed with caution and with appropriate disclaimers. Future research must identify reliable ways to measure app risk. The measure of app risks must be tested with end users to identify gaps in parameters used in the measure. The experiments presented here were conducted with participants online using web browsers and not with actual smartphones. Therefore, future research needs to explore the effect of risk communication and privacy priming on app choices on actual smartphones and in naturalistic settings over a long period of time.

ACKNOWLEDGMENTS

Removed for review.

REFERENCES

1. Mark S Ackerman and Lorrie Cranor. 1999. Privacy critics: UI components to safeguard users' privacy. In *CHI'99 Extended Abstracts on Human Factors in Computing Systems*. ACM, 258–259.

2. Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. 2015. Privacy and human behavior in the age of information. *Science* 347, 6221 (2015), 509–514.

3. Ralph Adolphs. 1999. Social cognition and the human brain. *Trends in cognitive sciences* 3, 12 (1999), 469–479.

4. Yuvraj Agarwal and Malcolm Hall. 2013. ProtectMyPrivacy: detecting and mitigating privacy leaks on iOS devices using crowdsourcing. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*. ACM, 97–110.

5. Steven Arzt, Siegfried Rasthofer, Christian Fritz, Eric Bodden, Alexandre Bartel, Jacques Klein, Yves Le Traon, Damien Octeau, and Patrick McDaniel. 2014. Flowdroid: Precise context, flow, field, object-sensitive and lifecycle-aware taint analysis for android apps. In *ACM SIGPLAN Notices*, Vol. 49. ACM, 259–269.

6. Farzaneh Asgharpour, Debin Liu, and L Jean Camp. 2007. Mental models of security risks. In *Financial Cryptography and Data Security*. Springer, 367–377.

7. Kathy Wain Yee Au, Yi Fan Zhou, Zhen Huang, Phillipa Gill, and David Lie. 2011. Short paper: a look at smartphone permission models. In *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices*. ACM, 63–68.

8. Kathy Wain Yee Au, Yi Fan Zhou, Zhen Huang, and David Lie. 2012. Pscout: analyzing the android permission specification. In *Proceedings of the 2012 ACM conference on Computer and communications security*. ACM, 217–228.

9. Kevin Benton, L Jean Camp, and Vaibhav Garg. 2013. Studying the effectiveness of android application permissions requests. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on*. IEEE, 291–296.

10. Alastair R Beresford, Andrew Rice, Nicholas Skehin, and Ripduman Sohan. 2011. Mockdroid: trading privacy for application functionality on smartphones. In *Proceedings of the 12th Workshop on Mobile Computing Systems and Applications*. ACM, 49–54.

11. Rainer Böhme and Jens Grossklags. 2011. The security cost of cheap user interaction. In *Proceedings of the 2011 workshop on New security paradigms workshop*. ACM, 67–82.

12. José Carlos Brustoloni and Ricardo Villamarín-Salomón. 2007. Improving security decisions with polymorphic and audited dialogs. In *Proceedings of the 3rd symposium on Usable privacy and security*. ACM, 76–85.

13. Jing Chen, Christopher S Gates, Ninghui Li, and Robert W Proctor. 2015. Influence of risk/safety information framing on android app-installation decisions. *Journal of Cognitive Engineering and Decision Making* 9, 2 (2015), 149–168.

14. Pern Hui Chia, Yusuke Yamamoto, and N Asokan. 2012. Is this app safe?: a large scale study on application permissions and risk signals. In *Proceedings of the 21st international conference on World Wide Web*. ACM, 311–320.

15. Eun Kyoung Choe, Jaeyeon Jung, Bongshin Lee, and Kristie Fisher. 2013. Nudging people away from privacy-invasive mobile apps through visual framing. In *Human-Computer Interaction–INTERACT 2013*. Springer, 74–91.

16. Vincent C Conzola and Michael S Wogalter. 2001. A communication–human information processing (C–HIP) approach to warning effectiveness in the workplace. *Journal of Risk Research* 4, 4 (2001), 309–322.

17. Nancy J Cooke and Steven M Shope. 2004. Designing a synthetic task environment. *Scaled worlds: Development, validation, and application* 263 (2004), 278.

18. Manuel Egele, Christopher Kruegel, Engin Kirda, and Giovanni Vigna. 2011. PiOS: Detecting Privacy Leaks in iOS Applications.. In *NDSS*.

19. Serge Egelman, Lorrie Faith Cranor, and Jason Hong. 2008. You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1065–1074.

20. William Enck, Peter Gilbert, Seungyeop Han, Vasant Tendulkar, Byung-Gon Chun, Landon P Cox, Jaeyeon Jung, Patrick McDaniel, and Anmol N Sheth. 2014. TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones. *ACM Transactions on Computer Systems (TOCS)* 32, 2 (2014), 5.

21. William Enck, Damien Octeau, Patrick McDaniel, and Swarat Chaudhuri. 2011. A Study of Android Application Security.. In *USENIX security symposium*, Vol. 2. 2.

22. Mica R Endsley. 1995. Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 37, 1 (1995), 32–64.

23. Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. 2012. Android permissions: User attention, comprehension, and behavior. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*. ACM, 3.

24. Vaibhav Garg and Jean Camp. 2012. End user perception of online risk under uncertainty. In *System Science (HICSS), 2012 45th Hawaii International Conference on*. IEEE, 3278–3287.

25. Ralph Gross and Alessandro Acquisti. 2005. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*. ACM, 71–80.

26. Jens Grossklags and Alessandro Acquisti. 2007. When 25 Cents is Too Much: An Experiment on Willingness-To-Sell and Willingness-To-Protect Personal Information.. In *WEIS*.

27. Sarah M Helfinstein, Jeanette A Mumford, and Russell A Poldrack. 2015. If all your friends jumped off a bridge: The effect of others actions on engagement in and recommendation of risky behaviors. *Journal of experimental psychology: general* 144, 1 (2015), 12.

28. M Hettig, E Kiss, JF Kassel, S Weber, M Harbach, and M Smith. 2013. Visualizing Risk by Example: Demonstrating Threats Arising From Android Apps. In *Symposium on Usable Privacy and Security (SOUPS)*.

29. Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. 2013. Privacy as part of the app decision-making process. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3393–3402.

30. Swapna Kolimi, Feng Zhu, and Sandra Carpenter. 2012. Contexts and sharing/not sharing private information. In *Proceedings of the 50th Annual Southeast Regional Conference*. ACM, 292–297.

31. David Kravets. 2014. Spyware executive arrested, allegedly marketed mobile app for' 'stalkers". In *arstechnica*. 279–298.

32. Jialiu Lin, Shahriyar Amini, Jason I Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. 2012. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 501–510.

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

34. Alexios Mylonas, Dimitris Gritzalis, Bill Tsoumas, and Theodore Apostolopoulos. 2013a. A qualitative metrics vector for the awareness of smartphone security users. In *Trust, Privacy, and Security in Digital Business*. Springer, 173–184.

35. Alexios Mylonas, Anastasia Kastania, and Dimitris Gritzalis. 2013b. Delegate the smartphone user? Security awareness in smartphone platforms. *Computers & Security* 34 (2013), 47–66.

36. Helen Nissenbaum. 1998. Protecting privacy in an information age: The problem of privacy in public. *Law and philosophy* 17, 5 (1998), 559–596.

37. Henry L Roediger and Jeffrey D Karpicke. 2006. The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science* 1, 3 (2006), 181–210.

38. Irina Shklovski, Scott D Mainwaring, Halla Hrund Skúladóttir, and Höskuldur Borgthorsson. 2014. Leakiness and creepiness in app space: Perceptions of privacy and mobile app use. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2347–2356.

39. Joshua Sunshine, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor. 2009. Crying Wolf: An Empirical Study of SSL Warning Effectiveness.. In *USENIX Security Symposium*. 399–416.

40. Jeff Sweat. 2000. Privacy paradox: Customers want control–and coupons. *Informationweek* 781, April (2000), 52.

41. Latanya Sweeney. 2002. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10, 05 (2002), 557–570.

42. Amos Tversky and Daniel Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty* 5, 4 (1992), 297–323.

43. Haidong Xia and José Carlos Brustoloni. 2005. Hardening web browsers against man-in-the-middle and eavesdropping attacks. In *Proceedings of the 14th international conference on World Wide Web*. ACM, 489–498.

44. Yajin Zhou, Xinwen Zhang, Xuxian Jiang, and Vincent W Freeh. 2011. Taming information-stealing smartphone applications (on android). In *Trust and Trustworthy Computing*. Springer, 93–107.