Using Personal Examples to Improve Risk Communication for Security and Privacy Decisions

Marian Harbach, Markus Hettig, Susanne Weber, Matthew Smith
Usable Security and Privacy Lab, Leibniz University Hannover, Germany
{harbach,smith}@dcsec.uni-hannover.de, {hettig, suey}@stud.uni-hannover.de

ABSTRACT
IT security systems often attempt to support users in taking a decision by communicating associated risks. However, a lack of efficacy as well as problems with habituation in such systems are well known issues. In this paper, we propose to leverage the rich set of personal data available on smartphones to communicate risks using personalized examples. Examples of private information that may be at risk can draw the users' attention to relevant information for a decision and also improve their response. We present two experiments that validate this approach in the context of Android app permissions. Private information that becomes accessible given certain permissions is displayed when a user wants to install an app, demonstrating the consequences this installation might have. We find that participants made more privacy-conscious choices when deciding which apps to install. Additionally, our results show that our approach causes a negative affect in participants, which makes them pay more attention.

Author Keywords
Android; Consequences; Examples; Permissions; Personalization; Privacy; Risks; Usable Security.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces

INTRODUCTION
Communicating the risk pertaining to certain actions is a long-standing problem in human-computer interactions. A large amount of previous research has looked at how to effectively communicate security risks in general [7], for example arising from insecure SSL connections [23] or phishing [11]. Many IT security systems use a decision dialogue to provide the user with information about potential risks or privacy implications. However, recent research has repeatedly demonstrated that such dialogues are often ineffective and quickly ignored [6, 10] or provide information that is hard to understand for the user [14]. In contrast, Rader et al. found that informal stories influence security behavior and thinking as they are being relayed from one user to the next [21]. These stories offer concrete examples of good or bad things that happened to people which a user can relate to. Blackwell et al. [3] previously posited that abstract information in software causes a gap between system designers and users.

As smartphones gain popularity, they also get more important in many peoples daily life, managing a large variety of personal information, including emails, pictures, call logs, and text messages. On Android, this information is protected from unauthorized access using permissions. The user gets to decide whether or not he or she agrees to the capabilities an app will have and which information on the phone will be accessible by that app after it has been installed. The most important prerequisite for such permission systems to be useful and secure in general is that the user can understand and decide which (sets of) permissions are okay to grant for an app and which might be harmful in a given context. Since apps can harvest and send out private information on the first launch, we believe that current trial-and-error app installation behavior is critical from a privacy perspective. Previous research of Felt et al. has shown that only 17 % of Android smartphone users are consciously aware of the specific permissions an app demands during installation [13].

To overcome the abstract nature of the existing dialogue, we show the user personalized, concrete examples of capabilities the app would get and which information it can access, using personal information present on the phone. Our goal is to raise awareness and make users more cautious while installing apps. For example, such a personalized example during app installation can say: "*If you install this app, it will be able to access and delete the following of your photos*", followed by a sample of the user's actual photos contained on the device. Thus, the communicated risks address a concrete, personal piece of information while listing a known cause as well as concrete consequences that a user can easily imagine. This can then allow users to judge whether or not a risk is acceptable or not: I may trust an anti-virus app to possibly delete some infected files in order to protect me from malware. Yet, I do not trust a random game to have access to text messages from my partner.

We emphasize that while users are probably aware of the private information they have on their devices and which apps they have installed, the individual relationship between an app's permissions and the concrete pieces of personal and private information that can then be accessed is never explicitly clarified. Our approach aims to enable users to make in-

CHI 2014, April 26–May 01, 2014, Toronto, ON, Canada

ACM 978-1-4503-2473-1/14/04..$15.00.
http://dx.doi.org/10.1145/2556288.2556978

formed decisions about the risks that apps pose to their private information by actually demonstrating which pieces of information can be accessed.

This paper makes the following main contributions to the HCI community:

- We explore the effect of personalized security decision dialogues on the Android app installation process, leveraging the rich set of personal information available on smartphones.
- We show how to design decision dialogues with personal examples and present a prototype of this approach.
- We provide and discuss the results of two user studies, demonstrating the effectiveness of our approach compared to the standard dialogue.

Our results shows that we are able to cause a considerable amount of users to rethink an app selection because of risks arising from permissions, even after an installation decision has already been made. Additionally, users who saw the modified permissions dialogue payed significantly more attention to an app's permissions in our experiments. We also found that a negative affect can increase attention. In the following, we will outline related work, introduce and motivate our concept before presenting the results, a discussion and future directions.

RELATED WORK

On the theoretical side, the HITL framework [8] and the C-HIP model [24], which HITL is based on, are general information transmission models which describe the process of sending an information through a channel to the human receiver. After receiving the concrete information, attention must be shifted to this information before the comprehension process starts and eventually an appropriate reaction follows. It has been shown [2, 25] that a personalization of decision dialogues can support this process.

According to De Paula et al. [9], the central problem of human interaction with IT security systems is that users should be able to make informed decisions without further help. Additionally, Bravo-Lillo et al. [4] state that such dialogues should "inform clearly about the consequences of the actions" and about the risks involved and emphasize that it is important to inform about whether an action is safe or not right before the user ultimately decides. Bravo-Lillo et al. [5] also investigated the behavior of novice users confronted with situations in which they should make security decisions. It was shown that those users are not aware of the sensitivity of their data and mostly started to worry after deciding to allow access. They found that novice users often act with a "let's-see-what-happens" attitude without thinking about possible consequences.

Blackwell et al. [3] discussed the influence of abstract representations in computational tasks and suggests to center a user's understanding of a system around task completion to overcome abstractness. Similarly, Inglesant et al. [16] and Karat et al. [17] find that the use of natural language to create access control policies is beneficial for their quality. Also, Raja et al. [22] found that users have improved mental models of firewall operation when using metaphors to embed security decisions into an application context.

Egelman [10] recently investigated Facebook Connect security decision dialogues. The central finding was that habituation caused most users to ignore the content of the dialogue, even though it was modified. Showing users personal information from their personal profiles, such as gender, name, or relationship status and sexual orientation, which would actually be accessed by the third-party website, did not improve the dialogue's efficacy in their study. In contrast to this work, the displayed information was already disclosed to at least one online service and the pieces of information cited in the dialogues were not very abstract or technical.

All of the above approaches argue for more concrete and graspable information in decision dialogues. Concrete examples of undisclosed private information have, however, not been used to highlight risks and possible consequences to support the decision demanded from the user. To the best of our knowledge, this is the first evaluation of personalized decision dialogues that leverages the large amount of personal information contained on smarpthones.

Permissions

A large body of work on Android permissions was compiled by Felt et al. (e. g., [12, 13]), investigating how permissions are used, how users perceive permissions, their attitudes towards potential risks as well as additional models applicable to ask for permission on smartphones. Most relevantly, they found that only very few users (3 %) actually understood which assets were protected by a given permission.

Pandita et al. [20] investigated to what extent a user's expectations of permissions match the actually requested set of permissions using natural language processing. They propose to automatically extract justifications for permissions from app descriptions and flag those apps where not all requested permissions are justified in the description. They postulate that this can help users to make informed decisions.

Recently, Kelley et al. [18] argued for the need of privacy as part of the app decision making process. They inserted an overview of the private information an app will be able to access into the overview page of Google's Play Store. Their results show that they were able to influence the user's decisions compared to the existing permissions display. Their goal was to make privacy part of the decision process by abstracting permissions into a summary table.

In contrast, we attempt to improve risk communication by making permission risks graspable and hence understandable. We want to enable users to take an ultimate decision if he or she is willing to trust a certain app and its developer with access to personal information. Using only the approach of Kelley et al., a user might just count the checkmarks in their Privacy Facts display and therefore make a privacy-aware choice. However, an app with only two permissions can already cause great harm if the app has malicious intent. We believe that more information is necessary to clarify the risks pertaining to certain permission sets than can be fitted on an

overview page. Our approach is hence complementary to the Privacy Facts display of Kelley and colleagues.

DESIGN

To leverage the power of personal examples as security decision dialogues for app installation, this information needs to be presented to the user in a concise and appealing fashion. Additionally, it is paramount that the permission dialogue is able to make the user question an app selection that was already made. This is necessary since the permissions will be displayed on a separate dialogue, which only becomes visible after a button labeled "Install" has already been pressed and therefore a choice for an app has been made. In this section, we will present the design of our modified permission dialogue for Google's Play Store and discuss its rationale. During the development of the UI, we ran pilot studies to test prototype efficacy.

Permissions Visualization

The current permission dialogue of Google's Play Store by default only shows a small number of the 79 permissions[1] which are deemed to be most important. The remaining permissions which an app requests can be displayed by unfolding a hidden panel. As Felt et al. already reported in 2011, the Android OS defines a large number of permissions of which many are rarely used [12]. Therefore, we wanted to choose a representative set of permissions for our evaluation.

To find the most common permissions, we crawled the 34,875 most popular apps on Google's Play Store in early 2013 and counted which permissions are requested. From the top 20 of the set of requested permissions, we picked ten (see Table 1) that can affect private information.

Permission	Rank #	Requested By
full network access	1	82 %
modify external storage	3	56 %
read phone status and identity	5	42 %
view Wi-Fi connections	8	26 %
precise location	9	23 %
find accounts on the device	12	16 %
take pictures and videos	14	8 %
read contacts	15	7 %
read call log	17	6 %
retrieve running apps	18	6 %

Table 1. The permissions selected for our evaluation, their rank in the top 20 of permissions, and how many of the 34,875 crawled apps requested them.

To visualize each of these, several random examples are selected from the data that this permission allows access to and displayed alongside a concrete, one-sentence description mentioning the user's actual data. Choosing random examples for each permission can prevent habituation, as different content is visible each time the dialogue is shown. Additionally, not every example drawn from the available information on the phone is considered equally private and displaying multiple examples shows the user a cross-section of the private information available on the phone.

[1]Google does not specify the actual number. According to Au et al. [1], Android 4.0 offers 79 permissions that can be requested by a third-party app.

Three of the permissions are notable exceptions: *full network access*, *take pictures and videos* and *location* do not allow access to existing data but can be used to exfiltrate information from the device or eavesdrop on the user's actions or surroundings. In these cases, we only used a description for the *full network access* permission or used the actual camera view for the *take pictures and videos* permission and the current location for the *location* permission.

We used the same general layout, fonts and headlines as in the existing permissions dialogue for all visualizations to match the look-and-feel of the original store. We slightly increased the size of the headlines for each permission from 18 dp (density-independent pixels) to 20 dp to make them stand out more clearly from the examples themselves.

Each of the permissions and its visualization are shown in Figure 1. We changed the descriptive text to mention real data and showed pieces of this data where possible. For example, for *full network access*, the user would be warned that this permission can be used to download malware to his or her phone or upload private information. Similarly, *modify external storage* showed a different picture from all the pictures taken with the phone and accessible via the storage permission each 1.5 seconds and stated that this app would be able to delete those. The phone's IMEI number and a statement that this number could be used for tracking or abused by malware were presented when the *read phone status and identity* permission was present. The remaining permissions gave examples of the information that could be accessed by this app alongside a statement saying "This app can see ..." or "This app has access to ...". The display would then show a selection of three contacts, three previous calls, three accounts configured on the device, three currently running apps, three nearby Wi-Fi networks and the current location on a Google Maps satellite image. In each case, we explicitly mentioned that the app to be installed will be able to interact with the concrete piece of data, in order to create a graspable connection between the app and the private information.

Overall, our permissions display was modified to include a different descriptive text more related to real data and a personal example of that personal data where possible. Additionally, the headline font size was slightly increased.

In a pilot study, we monitored participants' reactions to each of the permission visualizations and ordered them so that those deemed most relevant would be shown first. Contacts, call log and photos received the most attention and reactions from participants. More technical permissions such as access to the IMEI number and a list of Wi-Fi networks were therefore moved to the end of the list.

The remaining permissions were not part of our visualization but could be displayed as a fold-out panel at the bottom of our modified permissions display, similar to the existing approach. The choice of which permissions to display and which examples to use potentially influences the efficacy of our approach. We will discuss possible implications in the *General Discussion* section after presenting the results of our evaluation.

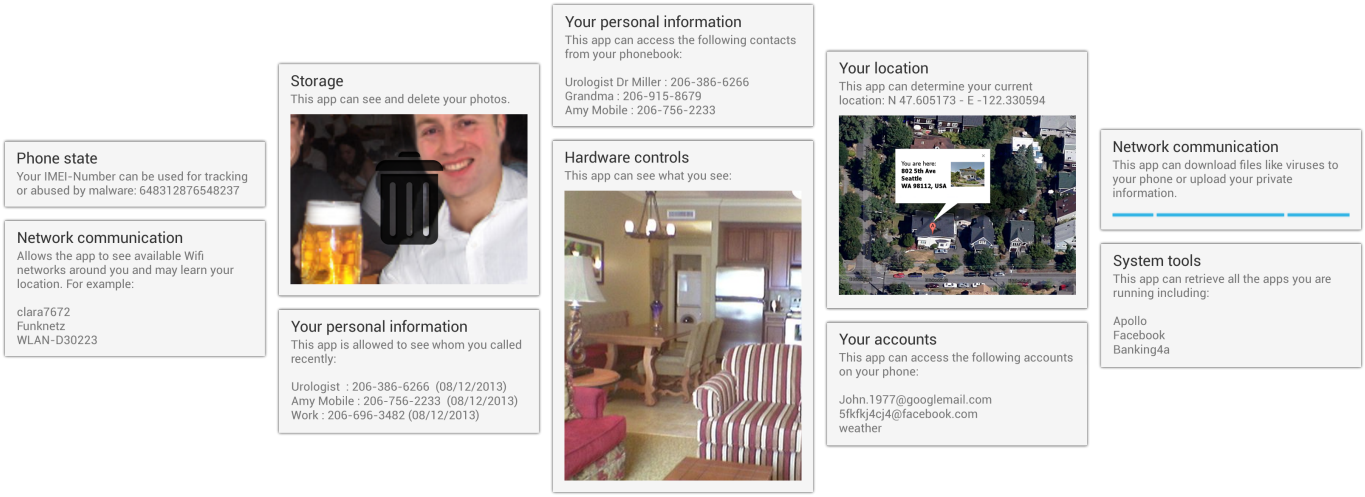

Figure 1. Overview of all permission visualizations created for our study.

Play Store Integration

As already argued above, we chose to replace the existing Play Store permissions dialogue, which acts as an ultimate decision before the actual installation. Figure 2 shows a comparison of the old and our modified permissions display. We created a working Android app simulating the existing Play Store, serving as a prototype for the evaluation.

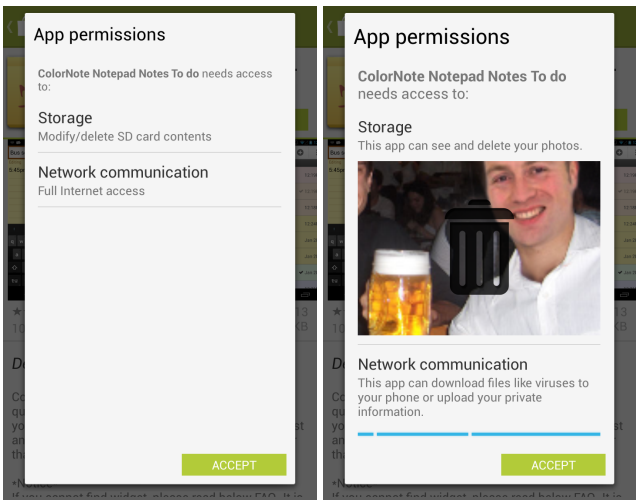

Figure 2. Comparison of the existing permissions dialogue on the left and our modified version on the right.

LAB STUDY

After piloting our prototype with several users, we set up a lab study to evaluate our approach and gain some further insights into how users select apps and perceive the risks pertaining to an app's permissions. Our methodology is based on the work of Kelley et al. [18]. To improve ecological validity, we let participants use their personal smartphones in this study and therefore base our evaluation on real private information.

Method

We invited owners of smartphones running Android 4.0 or newer from a university-wide study participation list to attend a lab study on app selection in the Google Play Store. We used limited deception by not mentioning the study's focus on permissions to prevent bias and recruited only Android users to prevent side effects from unfamiliarity with smartphones, apps or the Play Store. The invitation offered 8 Euros of compensation for a 30 minute study and stated that we would be installing a test app on users' personal device during the lab session. In the lab, each participant was introduced to the study and signed a consent form before installing our prototype app. Participants were informed that the app would be collecting usage data but no personal information and that we would transfer this information from their devices once the study was completed.

They were then instructed to complete six tasks, asking them to role-play the selection and installation of apps from a certain category that fit a certain purpose they were in need of. Participants were also asked to think aloud while making their decisions. Within each category, participants ultimately had to choose between installing one of two apps with different sets of permissions (see below). They were also instructed that they may choose not to install any app in each category, if none of the available options suited their personal needs. We added the "none" option to cater for the fact that users can normally abort an app choice process at any time without installing an app. In terms of study design, both, ours as well as the approach of Kelley et al. have drawbacks. Adding the "none" option allows participants to ignore a difficult choice, while forcing a decision may not represent realistic behavior. After finishing the tasks, participants completed a questionnaire on general app installation behavior as well as their perception of permissions. At the end, we debriefed participants about the true purpose of the study, clarified that our Play Store app was only a mockup and no apps were installed, removed all traces of the app from the participants' phones and gave them an opportunity to ask further questions.

The study used a between-subjects design with respect to the permissions dialogue and we compensated for effects of fatigue and learning by randomizing the task order. The task order assignment was based on latin squares.

Test App

To run this test, we implemented a mockup of the Google Play Store app as mentioned above. The mockup would either display the conventional representation of permissions as text or our representation using personal examples (cf. Figure 1). Before displaying the permissions, users were able to navigate through the Play Store as usual, compare several apps in the list view and look at app details, screenshots, and ratings. Each participant was provided with six tasks that asked them to find, select, and install an app with a certain purpose from a certain category of the Play Store. Our mockup app measured the time participants spent looking at each app description and the respective permission screens.

Apps

Each category contained two apps that actually fit the given purpose for the respective task. Within each category, we included padding apps to create a more realistic Play Store mockup. All apps and their names, descriptions, screenshots, permissions, and ratings were taken from the real Play Store and the 12 relevant apps (cf. Table 2) were selected to have similar functionality, average rating, and visual appeal. Similar to Kelley et al. [18], we also displayed one 2- or 3-star rating, one 4-star and one 5-star rating for each app. We selected apps that were likely to be unknown to our participants. We also included two app categories to test additional factors, rating and brand, to allow for assessing the influence of ratings and brand recognition on risk perception. Therefore, one photo app had a medium rating of 3.4 and another a high rating of 4.7. In another category, the well-known *Google Search* app was available with the *Quick Search* app. Concerning other properties the apps were again as similar as possible.

We also largely preserved the existing permissions for the apps. Yet, we wanted to have a larger variety of permission differences between the apps to see if there is a threshold of difference in terms of permission sets that is necessary for our approach to work. We therefore added or removed one or two permissions in four apps (cf. Table 2). Most notably, we made the Tetrity app not request any permissions and PicsArt, in addition to having a high rating, request all but one permission.

Mockup Store

To offer an experience as realistic as possible, the mockup store completely imitated the functionality of the real Play Store except for three differences. First, all app listings only included seven apps to not frustrate participants or have them spend too much time going through all available apps. Apps that did not fit the purpose of the task were not selectable. Second, there also was no search function, since our pilot study showed that participants would get frustrated searching for apps they already knew but were not included in this test. Instead, we instructed them to navigate using the categories.

App Name	Network Access	Ext. Storage	Phone State	WiFi Connections	Location	Find Accounts	Take Pictures	Read Contacts	Read Call Log	Running Apps	**Total**
EasyMoney		●	●				●				**3**
CWMoney	●	●	●		●		●	○			**6**
ColorNote	●	●									**2**
CatchNotes	●	●	●		●	●	●	●	●		**8**
Tetrity	⊖		⊖								**0**
Traris Deluxe	●		●	●	●						**4**
Weather	●			⊖	●						**2**
Eye in the Sky	●	●		●	●						**4**
PhotoEffects	●	●		●							**3**
PicsArt (rating)	●	●	●	●	●	●	●	○		●	**9**
Quick Search	●			●							**2**
Google (brand)	●	●		●	●	●		●	●		**7**

Table 2. The apps used in our studies and their respective permissions. A ○ indicates an added permission and a ⊖ indicates a removed permission. Two apps were part of a particular category (upper apps always requested less permissions) and participants were asked to choose one of those or none.

Third, the apps that suited the respective tasks would be randomly displayed as first or second item in the "Top Free" list to facilitate discovery and again make the tasks less frustrating without creating a bias for one app due to its position.

Participants and Results

After pilot testing our experimental setup, $n = 36$ participants completed the study. Sessions lasted between 15 and 30 minutes including the task, the questionnaire and the debriefing. Participant demographics can be found in Table 3. For the Westin index, we used the most recent set of question from Westin's 2001 Internet Privacy survey [19]. We did not find any significant differences in the measured data based on these demographic properties, including privacy inclination.

N	36	
age	19 – 30 years	
	median 23 years	
gender	12	female
	23	male
	1	N/A
IT experience	7	with professional or educational IT experience
	2	students of computer science
Westin Index	18	privacy fundamentalists
	17	privacy pragmatists
	1	privacy unconcerned
Incidents	17	previously victims of online dangers
	3	unsure

Table 3. Participants demographics for the lab study.

The left hand side of Table 4 details the results of our lab experiment, providing installation counts for each app. Across all six choices, more participants opted to install no app of the available two with the modified permissions dialogue when compared to the existing Android permissions. Yet, as shown in Table 4, the effect is only significant in half of the app choices. Also, in four cases participants tended to install the

less-requesting app or no app rather than the over-requesting app. The well-known brand and higher rating of two of the over-requesting apps did not diminish this effect. Similarly, we did not observe any effects for the Tetrity app, requesting no permissions. The PicsArt app, however, requesting almost all permissions, had the greatest reduction of installation counts (yet not quite significant). Participants became aware of the large number of permissions requested, even though this app's rating was considerably higher than for the alternative. This is also a notable difference to the results obtained using the Privacy Facts display of Kelley et al., where the rating was actually more important than the privacy facts.

Participants only spent an average of 3.1 seconds (median 1.0s) looking at the old permissions and 7.6 seconds (median 2.4s) on the modified version. While these values differ significantly between the two permission displays (repeated measures ANOVA across apps, permission dialogue as between-subjects factor, omnibus $F(1, 34) = 4.98$, two-tailed $p = .03$), the time spent on each permissions dialogue is still very brief.

Installation Behavior

We also asked participants about their installation behavior with free apps. 19 (52.8 %) stated that they usually look at several apps and then install one of them. An additional 11 (30.6 %) said they would install multiple apps, try them and then uninstall apps that did not suit their needs. 5 participants said that they would use a mix of the previous strategies. One participant stated to just install a number of apps and trying them without paying much attention to any descriptions. We then asked participants to rate how similar their usual selection behavior is to their behavior in the study. All but four participants selected 5 or more on a scale from *not similar at all* (1) to *very similar* (7). This suggests that the observed app choices are a suitable approximation of real behavior.

We also asked our participants to state how frequently they have not installed an app because of several factors (cf. Figure 3). Most participants said to not have installed an app because of the rating or the cost of an app, while the number of requested permissions or improper permissions only caused about a third of the participants to not install an app more than five times. This underlines the small amount of time users spend viewing the permission display. Additionally, most participants did not see many risks arising from malicious apps in the Play Store: they gave an average rating of 3.4 (median of 3) on a scale from "very low risk" (1) to "very high risk" (7). However, there was a slightly higher concern for the amount of danger for private information arising from smartphone apps in general: participants gave a mean and median rating of 5.0 on a scale from "no danger at all" (1) to "great danger" (7).

Qualitative Insights

Relevant comments users gave while thinking aloud during the app selection tasks were collected by the interviewer. Most interestingly, many participants indicated to trust Google to curate the Play Store or to trust the community of Android users to flag malicious apps. This confirms the slightly better rating for risks arising from Play Store apps

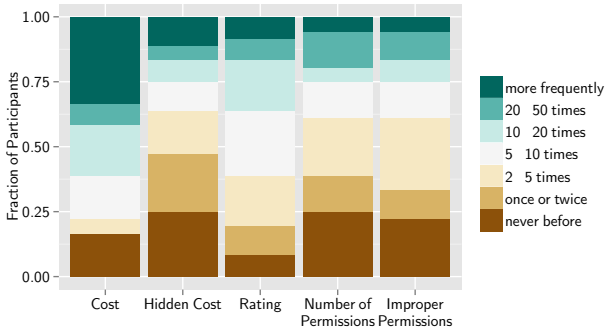

Figure 3. Frequency of not installing an app because of different factors from the lab study.

compared to apps in general presented in the previous section. Furthermore, several participants justified simply accepting any set of permissions without checking based on this view. Similarly, many participants stated to have "*nothing to hide*" or to be "*too uninteresting*" for anyone to attack them. Other participants also said that they do not put sensitive private information on their smartphone, because "*Google can see everything anyway*" or because existing protection mechanisms "*do not pose serious problems for hackers*".

We queried those participants who had seen the new permission dialogue about which of the private information contained in the dialogue was most sensitive. Those participants mostly referred to the pictures as being undesirable for others to have access to but also found that phone book, call log, and location can be sensitive. One participant said "*there are contacts in my phone book for which I would not like others to know that I am in contact with those*". Several participants hinted at a negative affect ("*I felt very alarmed which data can be accessed by apps*") or were very surprised how far access can go ("*I would not have thought that it can see or even delete all these things!*").

Discussion

The lab study yielded encouraging results. Our modified permission dialogue was able to generate a significant effect on the participants' installation behavior. The participants' comments support that seeing examples of their personal data encouraged reflection about possible consequences and whether or not they are ready to trust the developer and accept the risks. However, participants' comments and questionnaire responses also indicate that many other factors compete with permissions for the installation decision. A certain amount of the variation found in the results presented above is therefore accountable to users' different tastes in apps, as the baseline installs show. Several participants also stated to completely ignore permissions, because they don't feel that their personal data is threatened by apps or that they have nothing to hide. It was not our aim to make these users mind their privacy if they don't have an interest in it. However, we hope that concretizing the risks can influence some users' views of privacy.

The users' choices during the study also showed that, while we chose the four sets of apps with equal functionality, rat-

	Lab Study				**Online Study**			
App Name	Existing Permissions Dialogue (n=18)	Personalized Permissions Dialogue (n=18)	Δ_l	p-value	Existing Permissions Dialogue (n=157)	Personalized Permissions Dialogue (n=175)	Δ_o	p-value
EasyMoney	77.8 %	33.3 %	-44.5 %		36.3 %	33.7 %	-2.6 %	
CW Money	5.6 %	22.2 %	+16.6 %		49.7 %	28.6 %	**-21.1 %**	
none	16.7 %	44.4 %	+27.7 %	**.028**	14.0 %	37.7 %	**+23.7 %**	**< .0001**
ColorNotes	88.9 %	50.0 %	-38.9 %		79.0 %	57.7 %	-21.3 %	
CatchNotes	11.1 %	11.1 %	± 0 %		12.7 %	13.1 %	+.4 %	
none	0.0 %	38.9 %	+38.9 %	**.006**	8.3 %	29.1 %	**+20.8 %**	**< .0001**
Tetrity	61.1 %	55.6 %	-5.5 %		49.1 %	68.0 %	+18.9 %	
Traris Deluxe	38.9 %	27.8 %	-11.1 %		44.6 %	22.9 %	**-21.7 %**	
none	0.0 %	16.7 %	+16.7 %	.28	6.4 %	9.1 %	+2.7 %	**.0002**
Weather	66.7 %	77.8 %	+11.1 %		75.8 %	70.9 %	-4.9 %	
Eye in the Sky	27.8 %	5.6 %	-22.2 %		17.8 %	15.4 %	-2.4 %	
none	5.6 %	16.7 %	+11.1 %	.2	6.4 %	13.7 %	+7.3 %	.081
PhotoEffects	16.7 %	33.3 %	+16.6 %		19.8 %	28.0 %	+8.2 %	
PicsArt (rating)	83.3 %	50.0 %	-33.3 %		71.3 %	49.1 %	-22.2 %	
none	0.0 %	16.7 %	+16.7 %	.068	8.9 %	22.9 %	**+14.0** %	**< .0001**
Quick Search	27.8 %	27.8 %	±0.0 %		15.9 %	20.6 %	+4.7 %	
Google (brand)	72.2 %	44.4 %	-27.8 %		83.4 %	64.6 %	-18.8 %	
none	0.0 %	27.8 %	+27.8 %	**.041**	.6 %	14.9 %	**+14.3 %**	**<.0001**

Table 4. Installation count results of both, lab and online study. The table shows the percentage of participants that chose to install either one of two apps from each category or none, with the bottom app requesting more and unnecessary permissions than the top app within each group of two. Bold typesetting indicates significant contributions to the χ^2 value (standardized residual > 1.96) and significant p-values according to Fisher's exact test on the respective 3 (choice) $\times$ 2 (type of permission dialogue) cross-table.

ing and potential for visual appeal, some apps were already clearly preferred in the baseline condition. However, in the cases where the over-requesting app was preferred, the permissions dialogue still had an impact and was able to make users rethink their choice (cf., for example, the Weather or ColorNotes app in Table 4). Additionally, the effect we found on the rating-category underlines that we are indeed able to overlay decisions already made, even when some factors are strongly in favor of the more dangerous option.

Furthermore, the examples of private information displayed to communicate the risks of installing a particular app appeared to be making many participants understand the extent of what a permission allows an app to do. Even though all participants had been using Android smartphones for several months and had installed several new apps, seeing the permissions displayed in this new fashion gave them a better idea of what permissions really entail. Participants that came in contact with the new permission dialogue stated that they would "*pay more attention to these things in the future*".

ONLINE STUDY

To confirm the effects observed in our lab experiment on a limited population, we went on to evaluate our approach with a more diverse population. Since the lab study showed large individual differences in the way apps are selected and installed, a larger sample would be required to obtain more reliable results. Also, students are often considered to interact differently with technology than the general population and may have therefore biased our lab results. Additionally, we wanted to investigate to what extent our approach caused participants to be afraid of misuse of their private information. Such a negative affect could be a key motivation for less risky installation decisions. The questionnaire in this study therefore included questions that elicited how afraid of misuse of private information participants felt as well as how aware participants were of an app's control over personal data through the use of the permissions dialogue.

Method

We planned the online study to largely resemble the lab study. To access a diverse population, we decided to use Amazon's Mechanical Turk service to run the experiment. While MTurk also does not provide a sample that is representative of the general population, this service is commonly used to access a population that is more diverse than usually available. The task was framed as the common role-play of helping a good friend (previously applied in e. g. [5, 18]), who just bought a smartphone and does not know which apps to install. We showed participants a description of the scenario, including a picture of the friend John and his family to create a feeling of familiarity. Participants were asked to imagine that John gave them his phone and a list of activities he could use an app for. They were presented with the two suitable apps for each of the six activities. The order of the app choices was again randomized to control for effects of learning and fatigue. After finishing the selection process, participants completed a questionnaire as in the lab study and were compensated with $ 1.50. We asked participants to only participate if they use a phone running on Android 4.0 or newer in their daily life to have similar levels of familiarity and habituation with smartphones and app installations. We asked participants to specify their Android version in the questionnaire and excluded participants who specified a lower version.

As Amazon's TOS do not allow us to ask MTurk workers to install applications, we recreated the tasks in the browser.

Participants were presented with two side-by-side screenshots of the same apps for each of the same six tasks as in the lab study. Again, we chose a setup based on the work of Kelley et al. [18]. They were then asked to decide whether they would want to install any of the two apps. Again, they had the option to install none of the two. If participants chose to install an app, they were presented with a screenshot of the respective app's permissions dialogue (cf. Figure 2 and Table 2) and asked whether they want to continue with the installation or return to the app overview. Half of the participants were randomly assigned to our modified permission dialogue, which contained artificial private information of John (cf. Figure 2 and Figure 1). After each choice, participants were asked to explain the reasons for their selection.

Participants and Results

332 MTurk workers successfully and validly completed our task. We included attention check questions in the questionnaire and the tasks, which we used to exclude 49 participants answering inconsistently or specifying an Android version less than 4.0. Table 5 provides an overview of online study demographics. We found no differences in our measured variables with respect to these properties, except that participants indicating previous professional IT experience chose the Google app significantly more frequently (Fisher's exact test, $p = .035$, odds ratio .51).

N	332	
Age	18 – 64 years	
	median 27 years	
Gender	38.6 %	female
	61.4 %	male
Occupation	50.6 %	full-time employees
	19.3 %	students
	10.2 %	part-time workers
	7.9 %	self-employed
	7.0 %	Homemaker
	4.3 %	Unemployed or retired
IT Experience	27.1 %	have worked in or studied IT
Smartphone Use	18	months (median)
Westin Index	29.8 %	privacy fundamentalists
	52.1 %	privacy pragmatists
	18.1 %	privacy unconcerned
Incidents	38.9 %	previously victims of online dangers
	6.3 %	unsure

Table 5. Participants demographics for the online study.

During the app choice tasks, participants again were significantly less likely to install the over-requesting apps when using the permission dialogue with personalized examples as opposed to the baseline regular Android permission dialogue. The right hand side of Table 4 gives an overview of the results in comparison with the results from the lab. Again, brand and rating did not diminish the effect. The online study hence confirms the effect we found in the lab.

The Weather app was again so popular in the baseline condition, that we did not find significant effects in this case. However, this also suggests that we did not blindly scare users into not installing any app at all: the permission set of the Weather app was apparently reasonable enough to be accepted for its purpose. This is also mirrored in the participants' reasoning about this choice: "*The only permission it needed was my location. That makes sense.*" "*It need [sic] permissions which I would expect from this kind of app*". In other tasks, users responded differently: "*Both apps had permissions that I wasn't comfortable accepting*". Furthermore, we found that 88.6 % of online participants with the new permission dialogue installed three or more apps each.

Similar to the lab results, the permission dialogue had a significant overall effect on the time spent viewing the permissions (repeated-measures ANOVA across apps, permissions dialogue as between-subjects factor, $F(1, 330) = 53.98$, $p < .001$). Yet, Bonferroni-Holm-corrected pairwise comparisons only yielded significant results for the PicsArt and Quick Search Widget apps. However, the time spent looking at permissions was still short with 5.6 seconds (median 2.5s) in the baseline condition and 8.5 seconds (median 5.3s) in the modified condition. It is important to note that these values contain network and rendering delays, as they were collected on the server-side.

Again, we asked participants about their installation behavior with free apps. Similar to the lab, 208 (62.7 %) stated to examine several and then install one, 88 (26.5 %) try multiple apps and then uninstall those that did not suit their needs, and 31 (9.3 %) just try several apps without paying much attention to descriptions. 5 participants (1.5 %) said that they would use a mix of the previous strategies, while also considering external information sources as well as the required permissions. 89.4 % of participants again indicated to have behaved similarly in the online study, answering with 5 or more on a scale from *no similarity* (1) to *great similarity* (7). Online participants were also slightly more concerned about malicious apps in Google's Play Store (mean rating 3.96, median rating 4; 7=highest concern, see lab results) but slightly less concerned about the amount of danger for private information arising from smartphone apps (mean rating 4.2, median rating 4; 7=great danger).

Concerning participants' awareness of an app's control over personal data after completing the task, we found that the modified permissions dialogue had a significant effect: more than twice as many participants compared to the existing dialogue (87 vs. 40) chose the highest rating on a scale from *not aware* (1) to *very aware* (7) (odds ratio 2.88, Fisher's exact test, $p < .0001$). To assess how afraid people felt after seeing permissions, we asked them to rate to what extend the display of app permissions caused them to be afraid of John losing personal data or information. Similar to the general awareness, four times as many participants (13 vs. 57) indicated the highest rating on the provided scale from *not afraid at all* (1) to *very afraid* (7) (odds ratio 5.32, Fisher's exact test, $p < .0001$).

Most interestingly, we also found a significant effect of the above rating on the time spent viewing the permissions (repeated-measures ANOVA across apps, permissions dialogue and highest fear rating as between-subjects factor, $F(1, 328) = 124.3$, $p < .0001$). Participants that were very afraid to have John's private information compromised spent

5.4 seconds longer looking at both versions of the permissions dialogue. This is a twofold increase. Additionally, there was no significant interaction between the two between-subjects variables, such that we did not find differences in being afraid and paying more attention with regard to which permissions dialogue participants saw ($F(1, 328) = .007$, $p = .8$).

Discussion

Even though the setup of the online study differed from our lab study and did not allow us to use participants' actual private information to personalize the permission dialogue, the results show that the approach was still effective. Participants were more likely to choose an installation option that requested less permissions. Furthermore, the Weather app showed that reasonable sets of permissions were recognized and accepted if they fit the app's purpose. The online study also suggests that communicating risk to users with examples created more awareness in participants and instilled a negative affect which caused them to pay more attention to the permissions. We therefore believe that this approach to risk communication can more easily override choices already made than existing approaches, especially when important information on the risk can only be presented afterwards.

GENERAL DISCUSSION

The two studies have shown that we can leverage personalized, more concrete examples and descriptions in security decision dialogues to increase their efficacy. We were able to significantly impact the choice of apps in the Google Play Store, given that some apps were a greater risk to participants' privacy than others. The results also demonstrate that using personal examples in these dialogues can leverage users' affect to increase attention. While the modified dialogue may have caused some users to shy away from installing any apps in our study, it will also make them consider the privacy-tradeoff they are about to enter as it actually is. We might also be scaring users into choosing conservatively, but users were not completely oblivious to what happens, as in many cases, an app was still installed. Users that didnt install an app during the one try they had in the study, would probably try to find another one at a later point or when they have more options as they still have a need for the desired functionality.

For this exploration, we used a mix of drastic and more neutral scenarios in our permissions display. It is, for instance, rather unlikely that an app will simply delete images while it is more plausible for an app to access and possibly upload contacts from the address book. Since it is not possible to automatically determine what actions an app actually can or does take, it is up to future work to determine which intensity of descriptions and which visualizations can work best in which situations. We posit that a balance between intensity and frequency of displaying the dialogue needs to be found, since it is likely that if the most extreme examples are routinely used, they will lead to a fatigue effect if the examples shown never happen.

The concrete choice of examples is also likely to be an additional factor for the efficacy of the risk-communication-by-example approach. Displaying photos of one's cat or a random landscape from some city the person has visited probably causes less of an emotional reaction than a portrait of one's partner or child. Since we only displayed random examples in this initial exploration, selecting specific and more private examples can potentially further increase the efficacy of personalized security decision dialogues.

Furthermore, the comments users gave during the experiments indicated that using our prototype only for a short amount of time already caused them to consider changing their general attitude towards permissions ("*I think I will pay more attention to those permissions in the future.*"). Using personalized examples may therefore also be a suitable tool for an initial or recurring education campaign in systems that can serve to make users understand the risks present in this system.

The prototype we implemented for our evaluation assumes a single-user environment for the device. While this is reasonable in many cases, personalized decision dialogues may also be desirable in situations where a device is shared by multiple users. While a shared device, such as a tablet, will probably yield less information that is especially private to one of its users, a system would need to make sure that the personalized dialogues do not create a privacy issue themselves by disclosing private information to others. For mobile devices, this is an important and unsolved problem concerning many other aspects of the app model as well and is the subject of ongoing work (e. g., [15]).

Similarly, app installation may take place in public and hence cause privacy issues by displaying private information in the permission dialogue. To avoid this, users who frequently install apps in public should be able to opt out of using personal information or have the information displayed only after pressing a button.

Limitations

The study presented in this paper is a first contact study and thus the novelty of the dialogue itself may have increased the effects we observed. Additionally, we made the design decision to slightly increase headline size to separate the examples better, as well as to change the descriptive text of each permission to be less abstract and better fit the personalized context. These two changes may have confounded our results even though participants' qualitative reactions mainly concerned the personal information contained in the dialogues. Additionally, while we attempted to choose a set of apps which were not too well known, users may have already been familiar with some of the apps in our experiments, influencing their installation decision. Determining how each of these factors contributed to our results is the subject of future work.

As stated above, habituation can also occur with our personalized permission display. However, the displayed examples are chosen randomly and hence the dialogue changes between each app installation and therefore at least has a chance of countering habituation, especially since the users regularly create more personal data (e. g. new photos, contacts, and call log entries). This naturally needs to be investigated with a long term study before any reliable statements can be made.

CONCLUSION AND FUTURE WORK

In this paper, we have demonstrated the value of leveraging personalized examples to improve risk communication for security and privacy decisions, using the Android app installation process as an example. Two experiments with diverse populations have shown that users make more risk-aware app choices when presented with concrete examples of the information at risk from undesirable permissions. Furthermore, we found that when decision dialogues get personal, a negative affect was created in users which increased attention.

Thus, in future work, we would like to evaluate the long-term effects of this novel approach, especially with respect to habituation and general changes in behavior with respect to privacy. Since every interaction process uses a random selection of personalized information, our approach has the potential to counter habituation effects. Security decision dialogues with personalized examples can also be a valuable tool for other scenarios, including SSL warning messages, software installation on desktop computers or posting on social network sites. We also believe that personalized decision dialogues can serve as an educational tool, that may also work retrospectively. A user could be confronted with examples of private information, such as sent or received emails, that has been transmitted without proper privacy protection or encryption to see whether or not this causes more privacy risk awareness or a demand for better security measures.

REFERENCES

1. Au, K. W. Y., Zhou, Y. F., Huang, Z., and Lie, D. PScout: Analyzing the Android Permission Specification. In *Proc. CCS* (2012).
2. Besmer, A., Lipford, H. R., Shehab, M., and Cheek, G. Social Applications: Exploring A More Secure Framework. In *Proc. SOUPS* (2009).
3. Blackwell, A. F., Church, L., and Green, T. The Abstract is 'an Enemy'. In *Proc. Psychology of Programming Interest Group (PPIG) Workshop* (2008).
4. Bravo-Lillo, C., Cranor, L., Downs, J., and Komanduri, S. What Is Still Wrong With Security Warnings: A Mental Models Approach. In *Proc. SOUPS* (2010).
5. Bravo-Lillo, C., Cranor, L., Downs, J., and Komanduri, S. Bridging the Gap in Computer Security Warnings: A Mental Model Approach. *Security Privacy, IEEE 9*, 2 (2011), 18–26.
6. Bravo-Lillo, C., Cranor, L. F., Downs, J., Komanduri, S., Reeder, R. W., Schechter, S., and Sleeper, M. Your Attention Please: Designing Security-Decision UIs to Make Genuine Risks Harder to Ignore. In *Proc. SOUPS* (2013).
7. Bravo-Lillo, C., Cranor, L. F., Downs, J., Komanduri, S., and Sleeper, M. Improving Computer Security Dialogs. In *Proc. INTERACT* (2011).
8. Cranor, L. A Framework for Reasoning About the Human in the Loop. *UPSEC* (2008).
9. De Paula, R., Ding, X., Dourish, P., Nies, K., Pillet, B., Redmiles, D., Ren, J., Rode, J., et al. Two Experiences Designing for Effective Security. In *Proc. SOUPS* (2005).
10. Egelman, S. My Profile Is My Password, Verify Me! The Privacy/Convenience Tradeoff of Facebook Connect. In *Proc. CHI* (2013).
11. Egelman, S., Cranor, L. F., and Hong, J. You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings. In *Proc. CHI* (2008).
12. Felt, A. P., Chin, E., Hanna, S., and Wagner, D. Android Permissions Demystified. In *Proc. CCS* (2011).
13. Felt, A. P., Ha, E., Egelman, S., Haney, A., Chin, E., and Wagner, D. Android Permissions: User Attention, Comprehension, and Behavior. In *Proc. SOUPS* (2012).
14. Harbach, M., Fahl, S., Yakovleva, P., and Smith, M. Sorry, I Don't Get It: An Analysis of Warning Message Texts. In *Proc USEC* (2013).
15. Hayashi, E., Riva, O., Strauss, K., Brush, A. J. B., and Schechter, S. Goldilocks and the Two Mobile Devices: Going Beyond All-Or-Nothing Access to a Device's Applications. In *Proc. SOUPS* (2012).
16. Inglesant, P., Sasse, M. A., Chadwick, D., and Shi, L. L. Expressions of Expertness. In *Proc. SOUPS* (2008).
17. Karat, C.-M., Karat, J., Brodie, C., and Feng, J. Evaluating Interfaces for Privacy Policy Rule Authoring. In *Proc. CHI* (2006).
18. Kelley, P. G., Cranor, L. F., and Sadeh, N. Privacy as Part of the App Decision-Making Process. In *Proc. CHI* (2013).
19. Kumaraguru, P., and Cranor, L. F. Privacy indexes: A Survey of Westin's Studies. Tech. Rep. CMU-ISRI-5-138, Carnegie Mellon University, 2005.
20. Pandita, R., Xiao, X., Yang, W., Enck, W., and Xie, T. WHYPER: Towards Automating Risk Assessment of Mobile Applications. In *Proc. USENIX Security Symposium* (2013).
21. Rader, E., Wash, R., and Brooks, B. Stories as Informal Lessons About Security. In *Proc. SOUPS* (2012).
22. Raja, F., Hawkey, K., Hsu, S., Wang, K., and Beznosov, K. A Brick Wall, a Locked Door, and a Bandit: A Physical Security Metaphor for Firewall Warnings. In *Proc. SOUPS* (2011).
23. Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., and Cranor, L. F. Crying Wolf: An Empirical Study of SSL Warning Effectiveness. In *Proc. USENIX Security Symposium* (2009).
24. Wogalter, M., Dejoy, D. M., and Laughery, K. R. A Consolidated Communication-Human Information Processing (C-HIP) Model. *Warnings and Risk Communication* (1999), 15–23.
25. Wogalter, M. S., Racicot, B. M., Kalsher, M. J., and Noel Simpson, S. Personalization of Warning Signs: The Role of Perceived Relevance on Behavioral Compliance. *International Journal of Industrial Ergonomics 14*, 3 (1994).