Privacy as Part of the App Decision-Making Process

Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh

February 6, 2013

CMU-CyLab-13-003

CyLab
Carnegie Mellon University
Pittsburgh, PA 15213

Privacy as Part of the App Decision-Making Process

Patrick Gage Kelley
University of New Mexico
pgk@cs.unm.edu

Lorrie Faith Cranor
Carnegie Mellon University
lorrie@cs.cmu.edu

Norman Sadeh
Carnegie Mellon University
sadeh@cs.cmu.edu

ABSTRACT

Smartphones have unprecedented access to sensitive personal information. While users report having privacy concerns, they may not actively consider privacy while downloading apps from smartphone application marketplaces. Currently, Android users have only the Android permissions display, which appears after they have selected an app to download, to help them understand how applications access their information. We investigate how permissions and privacy could play a more active role in app-selection decisions. We designed a short "Privacy Facts" display, which we tested in a 20-participant lab study and a 366-participant online experiment. We found that by bringing privacy information to the user when they were making the decision and by presenting it in a clearer fashion, we could assist users in choosing applications that request fewer permissions.

Author Keywords

Privacy; Android; Mobile; Interface; Decision-making

ACM Classification Keywords

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

INTRODUCTION

In the past five years Android and iOS, the two now-largest smartphone operating systems, have transformed phones from devices with which to call others into true pocket computers. This has largely been accomplished through smartphone applications, often small, task-focused, executables that users can install on their phones from software markets. However, with each application a user downloads they may be sharing new types of information with additional app developers and third parties. Easy access to hundreds of thousands of applications from a diverse and global set of developers and the large amount of personal and sensitive data stored on smartphones multiply the privacy risks.

In Google Play, the current Android application marketplace, users are shown a series of "permissions" only after they have elected to download an application. Previous research suggests that users are likely to ignore the permissions display because it appears after they have decided to download a particular app [4, 12]. Furthermore, even users who pay attention to permissions displays have trouble using them because the screens are jargon-filled, provide confusing explanations, and lack explanations for why the data is collected.

Our research aims to provide an alternative permissions and privacy display that would better serve users. Specifically, we address the following research question: Can we affect users' selection decisions by adding permissions/privacy information to the main app screen?

To answer this question, we created a simplified privacy checklist that fits on the main application display screen. We then tested it in two studies: a 20-participant laboratory exercise and a 366-participant Mechanical Turk study. In each study we asked our participants to role-play selecting applications for a friend who has just gotten their first Android phone. Participants were assigned to use either our new privacy checklist or the current permissions display found in the Android market. Our results suggest that our privacy checklist display does affect users' app selection decisions, especially when they are choosing between otherwise similar apps. We also found that both the timing of the privacy information display and the content of the display may impact the extent to which users pay attention to the information.

Authors' preprint version
Accepted to CHI'13

To cite this paper:

Kelley, P.G., Cranor, L.F., and Sadeh, N.
Privacy as Part of the App Decision-Making Process. CHI 2013.

RELATED WORK

We outline previous research on the security model of the Android operating system, the current permissions model, and users' expectations regarding their phones. We focus on Android due to its historically more detailed permissions system and its large user base.

Android as a Major Application Provider

As of May 2012, Android has had over 15 billion application downloads, and over 500,000 applications, with both these numbers continuing to grow at an increasing rate [19].

Applications are not pre-screened for quality. Android app rating and recommendation site AppBrain reports that 33% of the applications in the Android Market are rated "low quality" by users. Additionally, a 2011 Juniper Networks report found "a 472% increase in Android malware samples" between July and November 2011 [11]. Similar studies from McAfee [16], Kaspersky Lab [20], and Symantec are all reporting continued exploits. The types and quality of this malware vary widely, ranging from attacks that collect user data (normally IMEI and other identifiers), to attacks that delete user data or send premium SMS messages.

To combat malicious applications Google internally developed a malware blocking tool codenamed Bouncer. Google announced that Bouncer had been checking "for malicious apps in Market for a while now," and as a result malware was

declining [18]. However, there are reports of Bouncer's limitations, such as applications existing in the market for weeks without being noticed [21].

Android Security Research

While Android has only existed publicly since 2008, a significant amount of work has been conducted on studying the Android permissions/security model. Much of this work focuses on creating theoretical formalizations of how Android security works or presents improvements to system security, and is largely out of scope. Enck et al.'s TaintDroid has bridged the gap between system security and user-facing permissions, focusing on analyzing which applications are requesting information through permissions and then sending that data off phone [5].

Vidas et al. also studied how applications request permissions, finding prevalent "permissions creep," due to "existing developer APIs [which] make it difficult for developers to align their permission requests with application functionality" [25]. Felt et al., in their Android Permissions Demystified paper, attempt to further explain permissions to developers [6]. However, neither of these papers explore end-users understanding of permissions.

There is also a growing body of work on the complexity of the current permissions schemes users must deal with. Researchers have discovered novel attack vectors for applications to make permission requests that are not reported to users [3]. Others who have looked at Android permissions have attempted to cluster applications that require similar permissions to simplify the current scheme [2] or have attempted a comparison of the differences between modern smartphone permission systems [1].

Android Permissions and Privacy Research

Android permissions are a system controlled by the Android OS to allow applications to request access to system functionality through an XML manifest. As these permissions are shown to the user at install time, this system as a whole forms a Computer-Supported Access Control (CSAC) system, as defined by Stevens and Wulf [24].

The majority of work done on user expectations related to this Android access control system has been done by our own group at Carnegie Mellon [12, 17] and two separate teams at Berkeley.

Felt and her colleagues have published a series of papers on the Android permission model, and how users understand it. They found that most users do not pay attention to the permissions screens at install time (83%) and that only three percent of their surveyed users had a good understanding of what the permissions were actually asking for access to [9]. They also performed a large risk-assessment survey of users' attitudes towards possible security and privacy risks, and possible consequences of permission abuses [8]. These results influenced our selection of items to include in a privacy checklist. Felt also performed work detailing other possible methods for asking for permission, with a set of guidelines for presenting these privacy and security decisions to users [7].

Moving away from permissions, the work of King et al. has explored user expectations across the entire use of their smartphones. This broader work, which included interviews with both iPhone and Android users, highlighted difficulties in recognizing the difference between applications and websites, personal risk assessments of possible privacy faults, and how users select applications in the application marketplaces [14].

Research in privacy policies, financial privacy notices, and access control have all similarly shown that privacy-related concepts and terms are often not well understood by users expected to make privacy decisions [13, 15, 22]. No work we are currently aware of has proposed and tested alternative permissions displays, or other ways to help users select applications in Google Play, or other application markets, as we do here.

PRIVACY INFORMATION IN THE ANDROID MARKET

This section details how Google Play currently presents privacy information and other information to consumers to help them select new applications to download to their Android smartphone. We then discuss the privacy facts display we designed to make privacy- and security-related information more central to users' selections.

Privacy currently in Google Play

Google Play users are presented with a number of ways to search and browse for new applications. Featured applications, top charts, categories, a search tool, and similar application lists each direct users to a common "Application Display Screen" (Figure 1A. Standard Market).

This screen provides users with a long list of information about each application. This includes (but is not limited to), a series of navigational items, application information, screenshots, a series of market-assigned labels (top developer, editor's choice), free-test descriptions, a series of reviews, and a series of other types of applications that users may have viewed or chosen. The current market application display screen is very long, yet completely lacks privacy information.

Privacy/security information appears on the above screens only when it is mentioned in free-form text by developers or when it appears in text reviews (almost always in a negative context). Market-provided (and by extension, system-verified) privacy/security information appears only on the secondary screen shown after a user has clicked the download button.

This secondary screen, where permissions are displayed (Figure 1, A. Standard Permissions), again displays the application name, icon, developer, and top developer status icon. This is followed by a very large accept button, which is followed (thus after the action target) by a list of grouped permissions. Only some permissions are shown initially, followed by a "See all" toggle that expands to display the remainder of the permissions an application requests. Each of these permission groups can be selected to see a pop-up window that contains the definitions for each of the permissions in the selected group. Because there may be several grouped

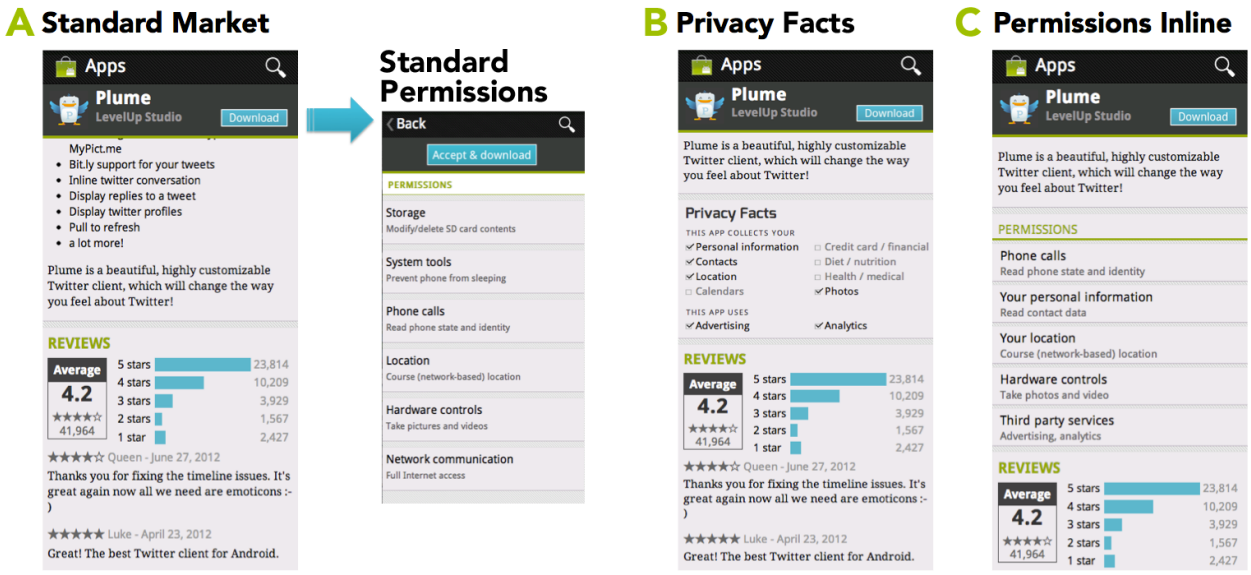

Figure 1. The three privacy/permissions display conditions we tested in our experiments.

permissions, the pop-ups may have to be scrolled to be read completely.

Reasons for modifying the Android application display

We posit that by the time a user selects to move forward by tapping the Download button, they have already made their purchase decision. We will see that this is true within our interview study below. For privacy information to be a salient part of the decision process, it must be presented to the user earlier in the process. Privacy information could be included in the long list of other application aspects on the standard application screen. Instead the current market places permissions on a secondary screen. While some might argue that placing permissions on their own screen draws users' attention to them, our results suggest that it actually does a disservice to users because they are unable to consider permissions as they consider other app characteristics.

Prototype privacy facts checklist design

We created a series of several possible locations and distinct styles of display in an ideation round. The custom privacy display that we decided to test is the Privacy Facts Checklist display shown in situ (Figure 1B. Privacy Facts). The display has several features:

Information— The display has two areas of information. The first with the header "THIS APP COLLECTS YOUR," describes eight types of information the app may collect: Personal information, contacts, location, calendars, credit card/financial, diet/nutrition, health/medical, and photos. The second header specifies "THIS APP USES" and lists advertising and analytics. Each of these ten items has a checkbox next to it, indicating use.

Display Style— The display is 270 pixels tall and the full width of the device (matching other standard application display sections). For comparison, the rating histogram is 162 pixels tall and the screenshots are the same as our privacy display at 270 pixels.[1] The display has a bold header "Privacy Facts" in a non-Android-standard type.[2] The remainder of the display is presented in the standard Android Market typeface. The items are each displayed at the standard size, with the headers in capital text in a lighter font color.

Location— The display is shown immediately after the Description section (and Video and What's New sections, if present, which they were not in our studies) and always immediately before the Reviews section. This means when participants first see each app screen there is no visual difference from the market as it is currently displayed, as the Privacy Facts section appears below the fold (as it would on most phone models).

Permission mapping— For this display we strayed from the current Android permissions by:

- Including types of information being collected that fall outside of the scope of the current permission model (health information, other financial information).
- Including the use of third-party modules, specifically advertising and analytics.
- Removing permissions that are nearly always used (Internet) and those that are irrelevant to most users such as networking protocols and rarely used permissions.
- Including photographs, which are currently accessible to applications.

The final selection of the checklist items we used was strongly influenced by the work of Felt et al. [8] as well as our earlier work [12], and a series of online pilots. The checklist includes

[1]There is variation in screenshot size on different Android phone models. The measurements above, and throughout this paper, are from a Google Nexus One.

[2]The font used is Exo from the Google Font Library.

both Android permissions as well as user-provided information. We wanted this display to include both, for a more holistic privacy summary. Also, by including an item like photos, we create a display that is more in line with users' expectations (which universal accessibility of photos is not). A more complex form of this display could include information that explains how these permissions are used, what they are used for, or how frequently they are used.

METHODOLOGY

We will discuss two phases of experiments: a 20-participant laboratory exercise and interview study, and an online 366-participant MTurk app comparison survey.

In our studies we ask participants to actively consider how and why they download applications in the market, complete our application selection task, and then discuss that experience. In both studies, the core of the experiment was an application selection task using different market designs that vary in how privacy information is presented.

Our study design was based on a similar study run by a team of researchers at Berkeley. The researchers had participants decide whether to install applications on a computer to see whether people read license agreements at install time. Their users evaluated the software tools as complete packages, based on brand, design, functionality, and also End User License Agreements [10]. Similarly, we seek to understand whether people read the permissions display or our updated privacy facts display when installing software on an Android smartphone, and whether we can manipulate their decisions through improved design and information.

Application selection task

The main task asked participants to select one application from each of six pairs of applications we presented in our "custom Android market." We presented two applications for each of the six categories (below). All of the applications we used were real applications that could be found and downloaded in the market. Their names, screenshots, descriptions, features, ratings, and reviews were all authentic. However, we picked most applications in the 1,000 to 10,000 download range, such that the applications would not have been seen or used by most participants. We displayed three text reviews per application, one 2- or 3-star, one 4-star, and one 5-star review.

In four of the comparisons we tested applications that were roughly equivalent (Twitter, document scanning, word game, nutrition app). In each of these four cases participants were presented with two applications with different permissions requests, detailed in Table 2. In each of these choices one of the applications requested less access to permissions and personal information (low-requesting v. high-requesting).

We also tested two special-case comparisons, to begin to explore the effects of rating and brand. In the flight-tracking comparison, we modified one of the applications (FlightTracker, low-requesting), to have an average rating of 3-stars. All of the other applications in all categories had 4-star average ratings. In the case of streaming music apps, we tested Spotify, a highly-known (shown in pre-tests) application with over 50 million downloads. Nearly all of our participants recognized this application.

Lab Study

To test the privacy facts display, and explore our research question, we conducted a series of semi-structured laboratory exercises in July 2012 with 20 participants. This was a between-subjects design. For the main application selection task ten participants saw the privacy facts checklist, and the other ten saw the current Android permissions display. We performed exploratory follow-up interviews seeking broad understanding of participants' interactions with their smartphones as well as diving deeply into issues surrounding the display of permissions, understanding of the terms in the checklist/permissions display, the safety of Google Play, and possible harms of information sharing.

We recruited participants through flyers and local Craigslist postings. Each candidate filled out a short pre-survey online before the exercise, which allowed us to confirm they used an Android-enabled smartphone. We performed the study in an on-campus lab and audio recorded the interviews. Participants were assigned randomly to conditions (without any balancing for gender, time-using android, technical knowledge, or age). They were paid $20 for successful completion of the interview, in the form of their choice of Target, Starbucks, or Barnes & Noble gift cards.

Exercise and Interview focus

The lab study followed a semi-structured format, outlined here:

- *Android introduction*: Questioned participants about general Android experience
- *General new smartphone advice*: Asked for advice to give to a hypothetical friend and new smartphone owner
- *Specific new smartphone advice*: Requested advice framed around a desire for six specific types of apps
- *Application selection task*: Had participants select applications with a Google Nexus One smartphone on our modified market
- *Post task explanation*: Requested explanations for why each app was selected
- *Android in the news and malicious activity*: Inquired on awareness of Android and apps in the news or on the Internet, then on malicious apps
- *Android permissions and privacy displays*: Drilled down to the privacy and permissions issues, asking if they had noticed the new display or used the current permissions display, depending on condition

Online Study

We conducted an online survey, a 366-participant MTurk test of the same application selection task used in the laboratory study. Because this was performed on MTurk the application selection task had a more structured survey format, as well as some other methodological differences that will be discussed

	Gender	Age	Occupation	Phone Model	Time Using Android	# of Apps Downloaded	# of Apps Frequently Used
P1	Female	21	Student	Motorola Droid	1–2 years	11–25	1–5
P2	Male	21	Student	Motorola Photon	7 months–1 year	101+	6–20
P3	Female	29	Other	Motorola Droid	1–2 years	11–25	6–20
P4	Female	39	Non-Profit	T-Mobile MyTouch 4G Slide	More than 2 years	11–25	20+
P5	Female	44	Marketing	Pantech Breakout Droid	7 months–1 year	11–25	6–20
P6	Female	30	Research / Science	Motorola Droid	1–2 years	11–25	6–20
P7	Male	43	Other	Motorola Droid	1–6 months	1–10	None
P8	Male	20	Student	Motorola Defy	1–2 years	1–10	6–20
P9	Male	31	Healthcare / Medical	Motorola Droid	More than 2 years	1–10	1–5
P10	Female	23	Research / Science	Samsung Galaxy	1–2 years	11–25	1–5
A1	Female	20	Student	Motorola Droid	7 months–1 year	11–25	1–5
A2	Female	23	Don't work	T-Mobile G2/HTC Desire Z	More than 2 years	1–10	1–5
A3	Female	20	Student	LG Ally / Optimus	1–2 years	26–100	6–20
A4	Female	28	Student	Samsung Galaxy	More than 2 years	11–25	6–20
A5	Female	24	Student	HTC rezound	More than 2 years	11–25	1–5
A6	Female	24	Research / Science	LG Ally / Optimus	1–2 years	11–25	1–5
A7	Male	21	Student	Motorola Droid	1–2 years	26–100	6–20
A8	Female	23	Research / Science	T-Mobile HTC G2	1–2 years	11–25	1–5
A9	Female	26	Research / Science	HTC Status	1–2 years	26–100	6–20
A10	Female	44	Healthcare / Medical	Motorola Droid	1–2 years	26–100	6–20

Table 1. Basic demographics of our lab study participants. Participant numbers beginning with P saw the privacy facts checklist, those with A saw the standard Android system. All the information above is self reported.

below in the limitations section. We again used a between-subjects design, but with three conditions. Participants saw one of: the privacy facts checklist (Figure 1B); the current android permissions display (Figure 1A); or the current android permissions display style and terms, presented in the application display screen with additional terms to cover categories from the privacy facts display (Figure 1C). In each case they were asked to pick six from the same 12 applications that our participants in the lab study were given, and then were asked to write a short sentence explaining their choice. For successful completion of the survey turkers were paid $0.30.

We used MTurk's user filtering system (95% success required) and required English speakers and Android users. The survey was front loaded with questions about the turker's Android device to discourage users who did not use Android phones. We manually inspected free-response questions to check for participants who were answering randomly, but removed no participants in that stage, only filtering (12) users who had not used the Android market.

LAB STUDY RESULTS

In this section we detail the results from our lab study. We cover the basic demographics of our participants, their experience with Android, their advice both general and specific to their hypothetical friend, the results of their application selection, and their post-task interview responses.

Demographics

As shown in Table 1, 25% of our 20 participants were male and 75% were female. Participants were between 20 and 44 years old, with an average of 28; 30% were undergraduates. All of our participants had downloaded Android applications from the market and were neutral or satisfied with the Google Play experience.

Application selection

The Privacy Facts display appears to have influenced participants in two of the four standard comparisons and in both of the special comparisons. Full selection percentages can be found in the first two columns of Table 3 (alongside the online study results).

In two of the four standard comparisons (word game, and Twitter) participants who saw the privacy facts display were, on average, more likely to pick the application that requested fewer permissions. In Document scanning, only one participant in each condition did not pick DroidScan Lite (the low-requesting app). In the diet application choice, no participants in the Android condition picked Doc's Diet Diary (the high-requesting app), while three with the Privacy Facts display did. In both the two special comparisons more of the participants who saw the privacy facts display picked the low-requesting app.[3]

Participants placed substantial weight on the design and perceived simplicity of using the application. Participants continued to surprise us with ever more idiosyncratic reasons for selecting certain applications. One participant preferred applications with simplistic names, saying "I like to download the apps that have a name that I can easily find. So Calorie Counter, I know where that is gonna be on my phone. I don't have to be like, oh, what is this called."

Participants reported wanting to try the apps out, often saying they would download many and see which was the best (which our study prevented them from doing). One said "And I might try things out and see... I just kind of see how well it works, because some things are more glitchy."

[3]Given the small numbers of participants we did not expect differences to be statistically significant and only the Twitter application choice was significant (Fisher's Exact test, $p = 0.023$, the odds ratio is 11.64).

	Personal	Contacts	Location	Calendars	Financial	Diet/nutrition	Health/medical	Photos	Advertising	Analytics	**Total**
Wordoid!	–	–	–	–	–	–	–	–	–	–	**0**
Word Weasel	✓	–	✓	–	–	–	–	–	–	✓	**3**
Twidroyd	✓	–	–	–	–	–	–	✓	–	–	**2**
Plume	✓	✓	✓	–	–	–	–	✓	✓	✓	**6**
DroidScan Lite	–	–	–	–	–	–	–	✓	–	–	**1**
M. Doc Scan Lite	✓	✓	–	–	–	–	–	✓	–	✓	**4**
Calorie Counter	✓	–	–	–	–	–	–	–	–	✓	**2**
Doc's Diet Diary	✓	✓	✓	–	–	✓	–	✓	–	✓	**6**
Rdio	✓	✓	–	–	–	–	–	–	–	–	**2**
Spotify *(brand)*	✓	✓	✓	–	✓	–	–	✓	✓	✓	**7**
Flight Tracker	✓	–	✓	–	–	–	–	–	–	–	**2**
iFlights *(rating)*	✓	–	✓	✓	✓	–	–	–	–	✓	**5**

Table 2. The boxes checked in the privacy facts checklist for each application are shown above. In each application category, one of the two applications requested access to fewer permissions (low-requesting always shown first).

Possible hidden costs also impacted application selection. Several participants noted that while the music streaming applications were free (as were all the applications we tested), they might have to purchase a subscription, or be unable to access certain functionality after a trial period ended. Participants generally wanted to avoid applications where features would expire or that would require later costs, but more importantly they expected the details of these arrangements to be extremely clear in the descriptions.

Android in the news and malicious activity

Most participants reported not seeing much about Android in the news, and most of what they did see being comparisons between Apple's iOS and Android. When we asked about reports of malicious apps, or apps doing unintended things, participants said they had not heard about this. Many believed that it could be hypothetically possible. One participant said "Like, I have wondered, oh could an app be a virus," another "I've heard about viruses, that they can actually shut your computer or phone down. Spyware."

Permissions and Privacy terminology

To test whether the terms we selected for the Privacy Facts display were understandable, we asked participants to explain what each term meant. While most were very clear, Personal Information and Analytics were the two that participants had the most trouble with. Personal Information answers were often too broad, encompassing things we did not intend. For example, one participant defined it as "That would mean like... interactions within the phone, Gmail, Messaging, Calling different people."

Participants generally preferred the checklist and its terminology. One participant said, "[Privacy Facts is] very straightforward to me. And that is something I noticed, I was thinking, Oh this is cool, is this what they are doing now. That is why I didn't say anything about it. I can immediately go: No, Yes, No, Yes."

Only two participants explicitly mentioned privacy information in their application selection decisions, both in the privacy facts checklist condition. One participant, said, "If this one is offering the same thing and they want less of your information, I would go with the one that wants less of your information." This comment shows her awareness of the privacy information, but also that the functionality must be matched between apps.

Task time and permission views

Overall, the entire laboratory exercise ranged from 29 minutes to 59 minutes (average 39:53). Participants spent between 3 minutes and 47 seconds to 25 minutes and 6 seconds on the application selection task. There was no statistically significant difference between conditions (two-tailed t-test, $p = 0.726$), although participants who saw the privacy facts checklist took on average 50 seconds more (11:40 v. 10:51) to complete the task.

Across all participants in the Android permissions condition, the permissions screen was used by participants for about half the selection decisions. Four participants decided which applications they would select without ever looking at any permissions screens. Another four participants looked at permissions for all the applications they selected. A6 looked at both Twitter applications permissions, but did not look at the permissions for either of the flight applications. A9 looked at only the permissions for the Twitter application she selected and no other applications.

Across all 31 permission screen views, participants spent between 1 and 11 seconds looking at the Android permissions display. On average they viewed the permissions display for 3.19 seconds (median 2 seconds), including page load time, a minuscule amount compared to time spent on the applications display screen.

ONLINE STUDY RESULTS

In our online study, the application selection task was conducted on MTurk through a participant's computer, not a smartphone. Participants saw the applications presented at smartphone size, side-by-side in iframes. Participants selected the application they thought was better for their friend, provided a short text reason, and then rated each of the two presented applications on the likelihood that they would personally acquire it.

With this study, we introduce a third condition, called Permissions Inline. This treatment was designed to separate the location of the privacy information from its format. It showed the standard Android Permissions Display, but positioned on the app display screen (where Privacy Facts is located) rather than in the standard location after the user tapped "Download." This condition tested whether it was only the existence of any privacy information on the application screen that changed behavior, or the checklist format and position.

We used the graphic design of the permissions display from the current Google Play store; however, we modified the labels to present the same information as our Privacy Facts display (including health, nutrition, advertising, and analytics).

	Lab Study		Online Study							
	Privacy Facts *(n=10)*	Android Display *(n=10)*	*Android Display (n=120)*	Privacy Facts *(n=123)*	Diff. from Android	p-value	Permissions Inline *(n=123)*	Diff. from android	p-value	Inline v. Facts
Wordoid!	60%	50%	40.8%	61.0%			49.6%			
Word Weasel	30%	50%	59.2%	39.0%	**20%**	**0.002**	50.4%	9%	0.198	0.095
Twidroyd	70%	20%	25.0%	52.9%			35.8%			
Plume	30%	80%	75.0%	47.2%	**28%**	**< 0.001**	64.2%	11%	0.051	**0.014**
DroidScan Lite	90%	90%	73.3%	60.2%			62.6%			
M. Doc Scan Lite	0%	10%	26.7%	39.8%	–13%	0.031	37.4%	–11%	0.076	0.784
Calorie Counter	70%	100%	55.8%	73.2%			73.2%			
Doc's Diet Diary	30%	0%	44.2%	26.8%	**17%**	**0.005**	26.8%	**17%**	**0.005**	1
Rdio	40%	30%	17.5%	28.5%			22.8%			
Spotify *(brand)*	60%	70%	82.5%	71.5%	**11%**	**0.048**	77.2%	5%	0.340	0.381
Flight Tracker	40%	20%	40.8%	35.0%			37.4%			
iFlights *(rating)*	50%	80%	59.2%	65.0%	–6%	0.358	62.6%	–3%	0.601	0.791

Table 3. Application selections in the laboratory and online studies. The application that requested access to fewer permissions (the privacy-protective choice) is always displayed on top. Statistics for the online study are comparisons to the base Android display. The right-most column shows the significance between the checklist and the inline permissions. Differences in bold, Fisher's Exact. Comparisons with the Android display were planned contrasts. The final comparison between the permissions inline and privacy facts display is Holm-corrected with an adjusted alpha of 0.01667.

An example of this is shown in Figure 1C.

Demographics

Of our 366 MTurk participants 59% were male and 41% were female (markedly different from our lab study). Our participants were between 18 and 63 years old, with an average of 28. All of our participants had experience downloading Android applications from the market (the 12 who did not were removed from this analysis).

Application selection

Overall the privacy facts display (changed format and position) had a stronger effect on participants application selections than only moving the permissions inline (changed position).

Privacy Facts display

In three of the four standard comparisons, significantly more privacy facts participants than Android participants chose the low-requesting app. Only for the document scanner did more participants in the standard Android condition choose the low-requesting app, and this difference was not significant.

For the Twitter choice, nearly three-quarters of the Android display participants chose Plume (high-requesting). One participant captured many of the common reasons for making this choice, reflecting, "Plume has 35,000 more reviews, which suggests to me that this is the more popular, more frequently used application. The description includes a list of everything you can do with the app and those all seem like useful features." However when presented with the privacy facts checklist, the two applications were selected at almost the same rate, with slightly more selecting Twidroyd. Here participants noted and cited the permissions information. One stated, "I picked the one that respects privacy more. The other gets too much personal info." Another participant wrote, "Plume collects too many personal facts."

For the special comparisons, rating and brand recognition outweighed privacy. However, even when one of the choices was a well-known brand privacy facts participants were significantly more likely than Android participants to select the relatively-unknown, low-requesting choice. For the flight tracking choice, more participants chose iFlights (high-requesting) over Flight Tracker. Although participants thought iFlights "sounds like an iPhone port," many believed it had a cleaner UI, but the top reason given was the rating difference. Flight Tracker's 3-stars seems to have outweighed all other factors. For the streaming music choice, Spotify (high-requesting) had much higher brand recognition (although again, both are real services). In the Android permissions display condition over half of the people (66/104) who selected Spotify explicitly stated that they had already heard it was very good or that they or friends use Spotify. One participant said "Spotify is pretty popular and I have never heard of Rdio." Spotify collected much more information than Rdio. but in this case we see that brand information trumps privacy concerns, though there is still a significant shift (11%) in favor of Rdio in the privacy facts condition.

Permissions Inline

As shown in in Table 3, the permissions inline display, while in the same place and often more space-consuming than the checklist, did not have as large an effect on users' decisions. In only one of the four standard comparisons, the nutrition application, was this change significant, and in most cases it underperformed the checklist display (significantly underperforming for twitter apps). This suggests that in addition to moving privacy information to the application display screen, it is important to present that information in a holistic, clear, and simple way if it is to impact users' app selections.

Free responses

Across the free-text responses for why applications were selected by participants in the Android display conditions, privacy was only mentioned by one participant, and permissions were mentioned by four others. Across the privacy facts checklist condition privacy was mentioned by 15 participants,

and permissions were mentioned by seven more. Information or info were mentioned by 49 people in the privacy facts checklist condition, but by only six participants using the Android display. Based on these responses privacy and personal information seem to have factored more strongly into the decisions of those who saw the privacy facts checklist.

Similar to our lab study, many participants, when directly asked, said they did not notice the privacy facts checklist. Of the 125 participants who were shown the privacy facts checklist, 49 (39.2%) reported in a free-text response having not noticed or paid any attention to the display. Both those people who did and those who did not notice the display provided reasons for why they ignored it, or believed it was not necessary:

- "I noticed the Privacy Facts but it really didn't influence me that much. I feel like with social networking it's so much easier to get contacts, photos, or information of someone."
- "It didn't influence my decision even though i noticed it. I tend to pay more attention to ratings and usefulness then anything else."
- "No, not really. It's not the most important factor. I don't keep a bunch of vital personal info on my phone, so no worries. I think people who do are really stupid."

There were also users who found the privacy facts display helpful and made their decisions based on it:

- "Yes. I believe the privacy information is helpful. It would only bother me if I saw something that didn't make sense for the app to use. However, I am not terribly concerned about privacy."
- "Yes. It only influenced me if it seemed to be the only thing to distinguish between the two apps."
- "Yeah, I always check that stuff. I want to know exactly what is happening to and with my data from that program when I use it. It was useful though I wish some apps would go into greater detail."

Participants who both used and didn't use the display still had misconceptions about companies, sharing information, and the market. Many assumed that all applications collect the same information. One participant who didn't look at the display said she did not because, "I assume they always say the same thing...."

Participants also continued to believe external forces protect them. One said, "Yes I saw the privacy facts. That didn't really affect my decision as companies are required to protect consumer's information and companies don't really wanna get sued for breach of security so I am not worried about all that." Another stated the continued belief that the market is internally well-regulated, "I think it is trustworthy, I would assume google play keeps a tight leash on that stuff."

Finally, one participant gave an answer that applies quite broadly, and mirrors work by Staddon et al. [23], "Yes, I noticed the privacy facts but it didn't effect [sic] my decision because I don't really know what the negative impacts of the information they obtain would be." Understanding the potential harm in allowing access to certain types of data remains difficult for consumers both in smartphones and other digital domains.

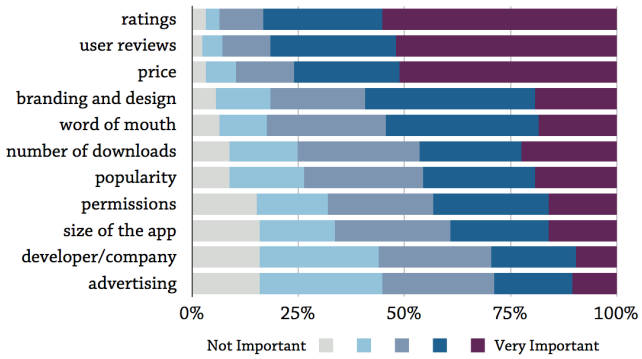

Figure 2. A series of factors users consider when decided on applications. Ranked by the number of users reporting a 4 or 5, where 5 is "Very important."

Self-reported decision factors

We also asked our online participants to rank a series of factors in their personal application-selection process from "Not important" to "Very important." The results of this are presented in Figure 2. Permissions ranked 8th (of 11), just below two metrics of popularity and just above the size of the application. 80% of participants said ratings and reviews were important or very important, compared with only 43% who said that permissions were important or very important.

This result seems to align with how often participants across our tests tended to ignore permissions.

Limitations

Our short checklist display had items that consumers were able to explain in most cases. Analytics and Personal Information were the most problematic. Participants were generally correct when defining Analytics, but often created more invasive definitions that were not intended. Personal Information was more difficult, as it was too vague and many participants listed other types of data that they then realized were covered by another item on the list. We will continue to further refine the terms and types of information that is most important to people.

One more significant design flaw with the display was that participants do not view permissions displays in the same way as they view privacy policies. They see this information only as items the phone can *take*, not things that they personally input. While we believe a complete privacy display should cover both user-provided information that is stored (i.e., medical or diet) *and* automatically collected information like location, this was not explained well by the current design.

Our lab study has many more female participants than male participants, and due to random condition assignment they were not evenly distributed. We note this as a potential limitation, though our results from the two studies are aligned, and we did not see such a similar gender imbalance in our online study.

Mechanical Turk also has its own set of limitations and biases, which we attempted to counter through a careful survey design. While we compared our two survey phases, they did not follow identical methodologies. Our lab study was more realistic, with users using actual cell phones, when on MTurk users saw the applications side by side, and could make direct and visual comparisons. While the reasoning and behavior given seems similar, it is possible that our online survey users had an easier time making decisions, not due to our improved permissions display, but due to the side-by-side display. The only evidence we have to counter this is the permissions-in-place display did not perform as well as the privacy facts display, implying that the side-by-side display alone is not responsible for all the improvements we saw.

Finally, we tested only 12 applications in the studies described above (and an additional 12 in early pilots). We picked applications that seemed similar, functional, and would be unrecognized, but we would like to expand this work in the future to consider larger application datasets.

DISCUSSION

Our goal was to better understand how users select Android applications, and to make privacy and permission information a salient part of that process.

We found that users did not use the current permissions display. By moving privacy/permissions information onto the main screen and presenting it clearly and simply we could affect user decisions in cases where applications were similar. Users mostly appreciated the new privacy facts display, said they would use it to help make their decisions or at least glance at it, and found comparing applications in the market to be a difficult task where better displays would assist them.

Can we affect users' decisions?

The short answer, is yes—the privacy information on the application display screen affected user behavior. In laboratory responses and our online test we saw behavioral differences as well as differences in quality and tone of responses relating to private information.

We also found most people do not consider permissions when downloading new applications. Even when instructed to download applications, most users made decisions without ever pushing the button that would take them to the permissions display. Both our lab participants and our online participants also self-reported that they were aware of the display, but did not look at it. This was confirmed by our lab study participants who, when they did fully "download" applications, spent a median time of 2 seconds on the permissions display. While this was expected based on other research and our own earlier work, we now have evidence that the permissions are, at least partially, disregarded due to their position in the application selection process.

In online testing we found that having a privacy display would in cases of equivalent applications change user application selection, and do so more strongly than simply moving the permissions display to the main screen. We also seem to have initial results that indicate even relatively sizable privacy differences cannot outweigh some other factors such as very popular applications (significant change, but not a majority) or differences in average ratings (3-star vs. 4-star apps, almost no change).

All of our participants had never seen a privacy facts display before, but were immediately able to make comparisons when specifically instructed to do so after the selection task. However, some simply did not believe privacy information was important or relevant to their decision. Some said it would depend on how much their friend (as part of the role-play) cared about his or her own privacy.

These results are similar to those seen in other labeling efforts. Consumers who care more about privacy, whether they have had a credit card stolen or have started receiving spam text-messages, are more likely to take advantage of labeling information. Even if the impact is not drastic, we see the privacy information on the main screen having an affect on selection behavior.

Do users enjoy, notice, and trust permissions information?

Participants in our studies reported being familiar with permissions displays and being aware that there are differences between applications. While this may seem unimportant or obvious, leveraging the awareness of privacy differences means creating interfaces, like checklists, that help consumers identify and compare differences should benefit users who want to make privacy-preserving decisions.

The terms on the current Android Permissions display remain difficult to understand and participants believed that there was little they could do as most of their information was already exposed. Participants reported that they did not, in most cases, read the information in the displays, and they did not select the permission groupings to see more details or try to better understand the terms. Even when the display was moved to the main screen, it does not have the impact of the privacy facts display.

Participants continued to report not being concerned with data sharing generally, partially due to a belief that companies are following laws and a strong belief that Android/Google is watching out for their safety as a consumer. While this is accurate in a very general sense, the specifics are quite far off from reality. Correcting the ubiquitous idea of Google Play as a safe, protected marketplace, must necessarily be changed if consumers are to protect themselves through understanding privacy and security in their decision-making process.

From both the lab and online studies we found that participants continued to report that other characteristics of applications are as important or more important than permissions, including: cost, functionality, design, simplicity, rating, number of ratings, reviews, downloads, size, and others. Continuing to understand how much privacy can compete and offset other aspects is important future work as consumers battle with a crowded and complex market.

When asked why an application was collecting a type of information, participants most often stated they did not know, but would occasionally venture possibilities. All of our lab

study participants wanted to better understand why applications required the permissions they did.

Finally, participants overwhelmingly trusted the application in both the privacy facts display and the permissions display. The question of trusting the information was one most had never considered, and actually gave some participants pause as they realized for the first time that this information might not be accurate. Again, users believe this information is correct, is being verified, and will assume they misunderstand something before they would believe the displays are incorrect. Mistakes in the permissions are not recognized, even when directly discussed. Users will assume they themselves are wrong, not the policy.

CONCLUSION

Smartphones have unprecedented access to sensitive personal information. While users are aware of this, generally, they may not be considering privacy when they select applications to download in the application marketplace. Currently, users have only the Android permissions displays to help them make these application selection decisions, screens which are placed after the main decision occurs, and are not easily understood. We sought to investigate how we could make permissions and privacy play a true part in these decisions. We created a short "Privacy Facts" display, which we then tested in 20 in-lab exercises and an online test of 366 participants. We found that bringing information to the user *when* they are making the decision and by *presenting* it in a clearer fashion, we can assist users in making more privacy-protecting decisions.

ACKNOWLEDGEMENTS

We acknowledge our colleagues at Carnegie Mellon University, and would like to thank Alessandro Acquisti and Sunny Consolvo. This work was supported by the National Science Foundation under Grant DGE-0903659 (IGERT: Usable Privacy and Security). Additional support was provided by NSF grants CNS-1012763, CNS-0905562 and DGE 0903659, by CyLab at Carnegie Mellon under grants DAAD19-02-1-0389 and W911NF-09-1-0273 from the ARO as well as Google.

REFERENCES

1. Au, K., Zhou, Y., Huang, Z., Gill, P., and Lie, D. Short paper: a look at smartphone permission models. In *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices (SPSM '11)* (2011).
2. Barrera, B., Kayacik, H., van Oorschot, P., and Somayaji, A. A methodology for empirical analysis of permission-based security models and its application to android. In *In Proceedings of the 17th ACM conference on Computer and communications security (CCS '10)* (2010).
3. Barrera, D., Clark, J., McCarney, D., and van Oorschot, P. C. Understanding and improving app installation security mechanisms through empirical analysis of android. In *2nd Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM)* (2012).
4. Egelman, S., Tsai, J., Cranor, L., and Acquisti, A. Timing is everything?: the effects of timing and placement of online privacy indicators. In *Proceedings of the 27th international conference on Human factors in computing systems*, ACM (2009), 319–328.
5. Enck, W., Gilbert, P., Chun, B., Cox, L., Jung, J., McDaniel, P., and Sheth, A. Taintdroid: an information-flow tracking system for realtime privacy monitoring on smartphones. In *In Proceedings of the 9th USENIX conference on Operating systems design and implementation (OSDI'10)* (2010).
6. Felt, A., Chin, E., Hanna, S., Song, D., and Wagner, D. Android permissions demystified. In *In Proceedings of the 18th ACM conference on Computer and communications security (CCS '11)* (2011).
7. Felt, A. P., Egelman, S., Finifter, M., Akhawe, D., and Wagner, D. How to ask for permission. In *USENIX Workshop on Hot Topics in Security (HotSec) 2012* (2012).
8. Felt, A. P., Egelman, S., and Wagner, D. I've got 99 problems, but vibration ain't one: A survey of smartphone users' concerns. In *2nd Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM)* (2012).
9. Felt, A. P., Ha, E., Egelman, S., Haney, A., Chin, E., and Wagner, D. Android permissions: User attention, comprehension, and behavior. In *Symposium on Usable Privacy and Security (SOUPS) 2012* (2012).
10. Good, N., Dhamija, R., Grossklags, J., Thaw, D., Aronowitz, S., Mulligan, D., and Konstan, J. Stopping spyware at the gate: A user study of privacy, notice and spyware. In *In Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS 05)* (2005).
11. Juniper Networks. Mobile malware development continues to rise, android leads the way, 2011. `http://globalthreatcenter.com/?p=2492`.
12. Kelley, P., Consolvo, S., Cranor, L., Jung, J., Sadeh, N., and Wetherall, D. A conundrum of permissions: Installing applications on an android smartphone. In *Financial Cryptography and Data Security*, vol. 7398. 2012, 68–79.
13. Kelley, P. G., Bresee, J., Cranor, L. F., and Reeder, R. W. A "Nutrition Label" for Privacy. In *Proceedings of the 2009 Symposium On Usable Privacy and Security (SOUPS)* (2009).
14. King, J. "How come i'm allowing strangers to go through my phone?"- Smartphones and privacy expectations, 2013. `http://jenking.net/mobile/`.
15. Kleimann Communication Group Inc. Evolution of a prototype financial privacy notice., February 2006. `http://www.ftc.gov/privacy/privacyinitiatives/ftcfinalreport060228.pdf`.

16. Labs, M. Mcafee threats report: Third quarter 2011, 2011. http://www.mcafee.com/us/resources/reports/rp-quarterly-threat-q3-2011.pdf.

17. Lin, J., Sadeh, N., Amini, S., Lindqvist, J., Hong, J. I., and Zhang, J. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. UbiComp '12, ACM (2012), 501–510.

18. Lockheimer, H. Android and security, 2012. http://googlemobile.blogspot.com/2012/02/android-and-security.html.

19. Lunden, I. Google play about to pass 15 billion app downloads? pssht! it did that weeks ago, 2012. http://techcrunch.com/2012/05/07/google-play-about-to-pass-15-billion-downloads-pssht-it-did-that-weeks-ago/.

20. Namestnikov, Y. It threat evolution: Q3 2011, 2011. http://www.securelist.com/en/analysis/204792201/IT_Threat_Evolution_Q3_2011.

21. Rashid, F. Y. Black hat: Researchers find way to "bounce" malware into google app store, 2012. http://www.scmagazine.com/black-hat-researchers-find-way-to-bounce-malware-into-google-app-store/article/252098/.

22. Smetters, D., and Good, N. How users use access control. In *In Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS 09)* (2009).

23. Staddon, J., Huffaker, D., Brown, L., and Sedley, A. Are privacy concerns a turn-off? engagement and privacy in social networks. In *Symposium on Usable Privacy and Security (SOUPS)* (2012).

24. Stevens, G., and Wulf, V. Computer-supported access control. *ACM Trans. Comput.-Hum. Interact. 16*, 3 (Sept. 2009), 12:1–12:26.

25. Vidas, T., Christin, N., and Cranor, L. F. Curbing android permission creep. In *W2SP 2011* (2011).

"I read my Twitter the next morning and was astonished" A conversational perspective on Twitter regrets

Manya Sleeper*, Justin Cranshaw*, Patrick Gage Kelley†, Blase Ur*, Alessandro Acquisti*, Lorrie Faith Cranor*, Norman Sadeh*

*Carnegie Mellon University
{msleeper, jcransh, bur, acquisti, lorrie, sadeh}@cmu.edu

†University of New Mexico
pgk@unm.edu

ABSTRACT

We present the results of an online survey of 1,221 Twitter users, comparing messages individuals regretted either saying in person (conversational regrets) or posting on Twitter. Participants generally reported similar types of regrets in person and on Twitter. In particular, they often regretted messages that were critical of others. However, regretted messages that were cathartic/expressive or revealed too much information were reported at a higher rate for Twitter. Regretted messages on Twitter also reached broader audiences. In addition, we found that participants who posted on Twitter became aware of, and tried to repair, regret more slowly than those reporting in-person regrets. From this comparison of Twitter and in-person regrets, we provide preliminary ideas for tools to help Twitter users avoid and cope with regret.

Author Keywords

Twitter; regrets; messaging; conversation; survey

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

It is easy to say something you regret, angrily insulting a loved one or inadvertently letting a secret slip. However, Twitter, a social networking service, enables these types of regrettable messages to spread rapidly and broadly, and to remain available for extended periods of time. Twitter's ability to broadcast messages widely and retain them indefinitely potentially alters the dynamics of regretted communications. In extreme cases, Twitter has enabled highly-publicized instances of regret, like Rep. Anthony Weiner's infamous tweet that led to his resignation [5]. However, everyday Twitter use can lead to more mundane regrets. As in conversation, Twitter users insult others, accidentally reveal private information, and express emotions in heated moments.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
CHI'13, April 27–May 2, 2013, Paris, France.
Copyright 2013 ACM 978-1-XXXX-XXXX-X/XX/XX...$10.00.

Thus it is worthwhile to try to understand regret both on Twitter and in in-person conversations. Past studies of in-person regret have identified factors that lead to regret, methods for becoming aware of regret, and strategies for repairing harm [8, 15, 16]. However, Twitter presents different features and limitations than offline conversation. Beyond offering wider audiences and increased message persistence, Twitter lacks face-to-face channels, such as body language, for transmitting apology or indicating offense.

We explore regretted messages Twitter users posted on Twitter or said during in-person conversations. We aim to improve understanding of regrets on Twitter by exploring similarities and differences with in-person regrets. By examining these regrets, as well as how people became aware of regrets in person and on Twitter, we also identify preliminary design directions for preventing and ameliorating regrets on Twitter.

Specifically, we examine four research questions:

- Q1: What states of being lead to regret on Twitter and in person?
- Q2: What types of regret occur on Twitter and in person?
- Q3: How do people become aware of regretted messages on Twitter versus in person?
- Q4: What repair strategies do people use to cope with regretted messages on Twitter and in person?

To address these questions we ran a 1,221-participant online Mechanical Turk survey with two samples. In one sample, we asked Twitter users to report on one message they regretted saying during an in-person conversation. In the other, we asked parallel questions about a message they regretted posting on Twitter. We collected information on the incident, the participant's emotional state preceding the incident, how the participant became aware of the regret, and any strategies used to ameliorate the regret. We used these answers to understand and compare drivers and consequences of regretted messages in in-person conversation and on Twitter.

BACKGROUND AND RELATED WORK

To demonstrate the conventions of posting on Twitter, we briefly highlight key features of the service. We then review related work on regret, first examining past analyses of regret during in-person conversations before discussing more recent work on social networking sites.

Twitter background

Twitter is an online social networking site where users post tweets, which are text-based messages of 140 characters or less. These messages are broadcast to a user's followers in relationships that are often asymmetric.

Twitter has several conventions that aid in sharing. Users can direct a message to a handful of specific users by crafting an @-reply. Users indicated by the @-reply will be alerted to the message through email or the Twitter client, but the message itself is public. A direct messages (DM) allows a user to send a private message to a single person. A users can also add #hashtags to a tweet to categorize it, better enable searches as part of a trend, or provide contextual information. Tweets are publicly accessible, unless an account is protected. Only a user's approved followers can view a protected user's tweets.

Related work

We first discuss research in the communications literature that has sought to understand many aspects of regret during in-person communication. We then outline work examining potentially regretful behaviors on social networking sites, as well as users' coping mechanisms.

Conversational regrets

Past studies of regrettable messages generally considered in-person messaging. Knapp et al. conducted 155 interviews, asking each participant about something they wished they hadn't said. The researchers noted eleven categories of regret, which we used in this study. They found that blunders, direct attacks, and group references were most frequently associated with regret, and that participants typically realized immediately when a message was regrettable [8].

Meyer studied individuals' cognitive states before and after messages regretted in person. She surveyed 173 undergraduates about their cognitive states prior to saying things they later regretted. Stress, frustration, and anger, as well as "having a lot on [their] mind," were most frequently associated with regrets. Participants commonly realized on their own that regrets had occurred, and regretted messages were rarely directed at more than one person [15]. Meyer separately examined efforts to repair the effects of in-person regretted messages. In a 204-participant survey she found that nearly two-thirds of repair strategies involved apologies, while excuses and justifications were also common [16]. We drew heavily from these studies for our survey design.

McLaughlin et al. also examined regretted in-person messages as part of broader work on failure events. They evaluated concession, excuse, justification, refusal, and silence as failure management strategies. They found that excuse was most commonly used overall, although increased guilt by the speaker tended to lead to concession [14]. We expect our work to more closely align with the narrower study of regretted messages.

Like in-person conversation, Twitter is primarily focused on individual short messages, potentially with intended audiences. Thus, to explore the regrets that emerge in Twitter communications, rather than in-person, we based our methodologies on the previous work and results from the in-person regrettable messaging literature.

Regret in social media

Regretted tweets have not been extensively studied; however, Wang et al. performed a mixed-methods study of Facebook users' most regretted posts. These posts often contained sensitive or potentially offensive content, were created during "hot," highly emotional states, or were seen by unintended audiences. The researchers also identified audience management and appearance management as major sources of potential regret [19]. While this work informed our analyses, differences between Facebook and Twitter usage patterns and audiences necessitated a different approach for investigating regrets on Twitter. Wang et al. looked at most regretted incidents, a method often used to examine life regrets. We instead looked at unspecified regrets, a method used for in-person messaging regrets.

Although regretted messages have not been directly addressed for Twitter, a variety of factors have been investigated that could contribute to, or help ameliorate, regret on Twitter. Marwick and boyd found that Twitter users deal with "context collapse." Users fashion tweets that can simultaneously fit a variety of social contexts by tweeting to an "imagined audience," employing self-censorship, or aiming to balance authenticity with conscious identity management [13]. This tension between perceived and actual audiences, as well as the difficulty of balancing authenticity with self-censorship, may lead to regret on Twitter.

Lampinen et al. found that users adopted proactive coping strategies for managing the co-presence of diverse groups on Facebook [10]. Wisniewski et al. also examined interpersonal boundary management on social networks and found that users adopted ad hoc boundary management mechanisms, such as ignoring information, blocking people, using aggressive behavior, or self-censoring [20].

Twitter presents a social environment with potential for regret. We seek to use methods from in-person messaging regrets work to understand the regrets that emerge, as well as the ad hoc awareness and repair strategies used to address such regrets, by examining and comparing Twitter users' regrets on Twitter and in conversation.

METHODOLOGY

Our goal was to analyze regrets that Twitter users had experienced on Twitter and during in-person conversations. We conducted a large-scale online survey from August to September 2012 using Amazon's Mechanical Turk (MTurk). We asked 1,221 MTurk Twitter users to each describe one thing they had said and then later regretted (the regretted message) either in an in-person conversation or on Twitter, depending on the sample to which the participant was assigned. We collected a description of the message, the context, how they became aware of the regret, and how they sought to repair the regret. It took participants 14.5 minutes on average to complete the survey, for which they were paid $0.75 (within the typical pay range for MTurk [6]).

Participant selection and samples

We screened for US MTurk workers over 18 years old who self-reported English proficiency and relatively frequent Twitter use (having had a Twitter account for at least a month and posting at least monthly). Of the 3,175 MTurk workers who started the survey, 946 were disqualified due to the age, language, or Twitter use requirements. The majority (609, 64.4%) were disqualified for posting less than once a month on average.

Survey

Samples

After the initial screening questions, participants were split into two samples in a round-robin fashion. The first sample was conversational regret, which mirrored previously described work. The second sample asked parallel questions, slightly reworded to focus on Twitter regret. In both samples, participants were asked to recall a time when they said or tweeted something and then regretted it, with the wording and format of the prompt based on work on in-person messaging regrets by Meyer [15, 16].

Our prompt for **conversational regret** participants was:

> "Please recall an occasion when you **said** something during an **in-person** conversation and then regretted saying it. This may be something that you regretted saying immediately or that you regretted saying later."

Our **Twitter regret** prompt was similar:

> "Please recall an occasion when you **tweeted** something and then regretted tweeting it. This may be something that you regretted tweeting immediately or that you regretted tweeting later."

Survey structure

Participants in both samples who could not recall a regret were directed to an alternate survey that asked them about why they did not have regrets. We do not report the results of this survey, as the goal was only to ensure an equal workload for either a positive or negative response. Of the 1,889 participants who qualified for the study, 601 (456 for Twitter and 145 for conversational regret) could not recall a regret.

Participants who were able to recall a regret were then asked to fill out a survey about the regretted messages they reported in response to the initial prompt. The survey drew heavily on questions and structure from in-person messaging regrets work [8, 15, 16] and included several groups of related questions. We asked participants to answer:

Regretted message description: a series of essay questions that asked the participant to describe the message in detail, including the context, the reason why they said/tweeted it, the intended audience, the audience's reaction, why they regretted the message and any consequences

Circumstances: follow-up questions about their state when they delivered the message

Awareness: free response about how they became aware that they should not have said the message, followed by a multiple choice selection of how quickly after the message they realized they should regret the message

Repair strategies: a description of whether, how, and how successfully they tried to repair any harm caused by the message; participants were also asked to rate the seriousness of the regret before and after repair

Twitter specifics: questions on Twitter usage (e.g, client and device tweeted from, is/was the account protected)

Demographics: basic demographic questions

We based the general survey structure on the format used for in-person regrettable messaging work [15, 16]. Specifically, we used Meyer's format of asking participants to provide one regret and then probing for details. Although this format has several weaknesses, as outlined in Limitations, it has been used repeatedly to examine in-person regrets.

Quality control on Mechanical Turk

While MTurk has been shown to produce quality samples and results [6], surveys on MTurk should be designed to encourage quality responses. We took several quality control measures. First, we only used MTurk workers who had over a 95% approval rating on the site. Second, we front-loaded longer essay questions. By putting these questions earlier in the study, we encouraged lazy or unmotivated participants to drop out early or to enter nonsensical data where it was visible. It also made it easy for honest survey participants to return to the task, without feeling like they still needed to invest large amounts of time. We removed a small number of participants (25) from the dataset who provided nonsensical or non-English answers to the free response fields.

We also removed responses from a small number of conversational regret participants (32) who responded about a regret on Twitter. We believe they did so because they were primed to think about Twitter when recruited as Twitter users. An additional 350 participants were removed for not completing the survey.

Data analysis

We surveyed MTurk users who posted on Twitter about a regretted message either said in-person or posted on Twitter. Although the surveys for each sample were designed to be parallel, the fundamentally different contexts preclude statistical comparisons between samples. To explore characteristics of how regret on Twitter compares with in-person regret, we present the results of the Twitter and conversational regret samples side-by-side. However, the proportion of participants reporting different answers is only meant to illuminate general themes and trends, not intended to be compared statistically.

Within a single sample, we perform statistical analyses. We use logistic regression to evaluate the relationship between types of regret and whether the audience was a group or individual, the relationship between awareness mechanisms and whether or not regret was experienced immediately, and the impact of the repair strategy on the success of reconciliation. Demographics were compared between samples using a

Wilcoxon test for numerical data and χ^2 tests for categorical data. All tests use a significance level of $\alpha = .05$.

Participant demographics

After quality control removals, 1,221 people reported regrets: 747 for conversational (79% of those who started) and 474 on Twitter (51%). The age and gender distributions were slightly skewed toward young and female participants. The mean age was 30.3 (28.2 for Twitter and 31.7 for conversational regrets). Overall, 53% of participants were female and 46% were male (10 preferred not to answer). The gender breakdown was almost identical for the Twitter and conversational regret samples. Approximately a quarter of participants were students (26%) and 10% were unemployed. The remainder were primarily employed in science (9%), service (8%), and art (8%) fields. There were no significant differences between the Twitter and conversational regret samples in age, gender or occupation, nor were there significant demographic differences between participants who did and did not report regrets.

ANALYSIS AND RESULTS

Q1: States of being leading to regret

States leading to regret

People often say things they later regret because of demands on mental capacity that impair thought processes. We found that both Twitter and conversational regret participants were often in negative, highly emotional states prior to regret. Meyer outlines several factors that contribute to "cognitive load," "physiological state," and "emotional state" that can potentially lead to regret [16]. We asked participants about these states and additionally, based on results from Wang et al. [19], asked participants whether they were drunk at the time of the message. For each, we asked participants to rate how much or how little it applied immediately before they tweeted or spoke on a five point scale, where a one indicated "Not at all" and a five indicated "Very much so." They rated each of the following: "I was fearful or frightened," "I had a lot on my mind," "I was feeling excited," "I felt ill," "I was worried," "I was nervous or anxious," "I was drunk," "I was angry," "I was stressed," "I was tired/fatigued," "I was happy," "I was hung over," and "I felt frustrated."

Consistent with Wang et al.'s work on Facebook regrets [19], we found that both in person and on Twitter, negative hot, or highly emotional, states emerged as most common prior to regret. Both conversational and Twitter regret participants commonly reported a four or a five for stress (46% for Twitter and 50% for conversation), anger (51% and 43%), or frustration (58% and 53%) prior to the regrets. Participants also often had something on their minds (54% and 51%). Somewhat less common were positive emotions including feeling excited (26% and 17%) or happy (22% and 21%).

Q2: Types of regret

We also looked at types of regrets participants reported for Twitter and for in-person conversations. For both, participants most commonly reported regrettable messages that were critical of others. However, on Twitter, participants more commonly regretted content that was expressive/cathartic and that was intended for groups of people.

Participant-reported Types of Regret

	Twitter		Conversation	
Reveal too much	117	25%	105	14%
Direct criticism	96	20%	213	29%
Expressive	64	14%	15	2%
Direct attack	62	13%	108	14%
Blunder	51	11%	120	16%
Implied criticism	34	7%	84	11%
Group reference	13	3%	21	3%
Agreement changed	3	1%	10	1%
Behavior edict	2	0%	28	4%
Lie	1	0%	25	3%
Other	31	7%	18	2%

Table 1. Types of regret for Twitter and Conversation

Types of regret

We coded each regret described by participants into one of Knapp et al.'s categories for types of regretted conversational messages [8], specifically:

- **Blunder:** "not normally perceived by a third-party observer as problematic"; mistakes, factual issues; includes typos or errors during conversation
- **Direct attack:** "critical statements directed at a person, the person's family, or the person's friends [...] general rather than specific"
- **Group reference:** stereotypical references about a group (e.g., ethnic, racial)
- **Direct criticism:** critical statements about "something specific" about a person
- **Reveal/explain too much:** telling "more than the situation calls for;" e.g., undesired personal information or a secret
- **Agreement changed:** agreeing to something, then later changing one's mind
- **Expressive/catharsis:** general "expressions of feeling and emotion"
- **Lie:** "knowingly lying to another person"
- **Implied criticism:** "critical remarks that are implicit" and can be "teasing remarks"
- **Behavioral edict:** telling someone to behave in a certain way

Two coders independently coded all the regrets based on Knapp et al.'s categories. Two of three coders reached a consensus for any regrets for which there were discrepancies.

Across both in-person and Twitter regrets, participants most commonly regretted critical statements (Table 1). Common critical statements included direct attacks and direct criticism; 29% of conversational and 20% of Twitter regrets were direct criticisms while 14% of conversational and 13% of Twitter regrets were direct attacks.

Blunders also arose frequently for both conversational and Twitter regrets, although more often for conversational (11% for Twitter, versus 16% for conversational). Although both Twitter and conversational regret participants reported some similar blunders, such as saying/posting messages they later found out were false or that had been said/shown to someone who found them offensive, some blunders were unique to Twitter. On Twitter, time-delayed blunders sometimes caused participants to regret messages because of an event or change in context. For example, one participant regretted tweeting about a drive-by shooting in his friend's hometown when that friend was later killed in a drive-by shooting. Twitter, as an online interface, also allowed blunders caused by typos and broken links, which several participants found embarrassing. For example, one participant reported being "made fun of" for tweeting that he "used a lot of hags on [his] car."

Participants also regretted expressive or cathartic content more frequently on Twitter than in person (14% versus 2%). These expressive statements were typically tweeted when participants were angry or upset and often served to vent or express frustration on topics such as work, relationships, or politics. Often, the goal was to allow other people to see and sympathize or "know what [the participant] was going through." Participants tended to regret the message later after re-thinking how it would sound, or after someone who viewed it became upset. For example, one participant described tweeting "Last day of my internship, so excited to be done," because she "was unhappy with how the internship treated [her] and what had happened [...and] wanted [her] friends to see it because they knew [she] was having a rough time." However, she regretted the tweet when her internship coordinators saw it and sent her an email telling her she needed to delete the tweet. In contrast, expressive regrets in conversations tended to be part of arguments or opinions.

Type and audience

Participants also specified whether they intended the messages to be seen or heard by individuals, or by multiple people. Twitter regret participants were more likely to target multiple people (73% of Twitter regrets, versus 24% of conversational), likely because of Twitter's broadcast capabilities.

Some types of regretted messages tended to be intended for multiple people more often, especially on Twitter (analyses using logistic regression). When their intended audiences were multiple people, rather than individuals, Twitter regret participants were significantly more likely to report a blunder ($p = 0.008$), content that revealed too much ($p = 0.005$), or expressive/cathartic content ($p = 0.003$). Of Twitter blunders, 82% were intended for multiple people, versus 33% of reported conversational blunders. Twitter regret participants often said that they wanted to tweet to friends, coworkers, or others interested in a specific topic, but regretted the tweet because they made an error that caused confusion or made them look bad. For example, one participant reported tweeting, "Congratulations to B for being elected ALA Councilor," intending the message for other librarians in South Carolina. She later realized that the individual was actually a candidate for the position, rather than having been elected and regretted the tweet because "it was embarrassing."

Twitter regret participants who regretted expressive or cathartic posts also tended to target multiple people rather than individuals (84% of expressive/cathartic regrets). Participants tended to want to share political or negative feelings with the general public or their friends because they "wanted to vent" or express their feelings "to anyone that would listen."

Regretted statements on Twitter that revealed too much also tended to be targeted at multiple people (80%). Many participants tweeted personal information, such as details about their lives or relationships, and then regretted sharing them on Twitter. Several participants also reported having both personal and professional accounts and regretting tweeting personal information on their professional Twitter accounts. For example, one participant said that he regretted tweeting "on my professional twitter account about a night of heavy drinking" because it seemed "unprofessional."

In contrast, conversational regret participants were significantly more likely to report regrets that were direct attacks ($p = 0.024$) when the intended audiences were individuals (67%) than multiple people. Participants were typically angry or arguing with the individuals and said the regretted message to them. For example, one participant "screamed at my father that 'I hate him' in an argument" because his father kept him from attending a party. On Twitter, such attacks were commonly focused at groups (68%), and participants reported wanting their anger to be seen by people. For example, one participant had a conflict with a friend, and wrote "she's so annoying and whiny," intending "it to be seen by friends."

Unintended audience

We also coded unintended audiences for the regretted messages. In conversation, unintended audiences included people overhearing messages (e.g., by walking into a room) or being told about them. On Twitter, most of the tweets reported were public tweets. However, participants still had particular audiences in mind when they tweeted. Unintended audiences occurred because people other than the intended audiences saw or heard about the tweets.

For Twitter regrets, 13% had unintended audiences, compared to 5% of in-person regrets. Unintended audiences occurred most commonly on Twitter for regrets that revealed too much (23% of regrets that revealed too much), often because participants tweeted something private, insulting, or about work, that they later realized they didn't want everyone to know. For example, one participant described how she tweeted "something sexual and my twitter at the time was public, so I freaked out when I saw that my brother's screen name popped up on Recommended Twitter."

Level of regret

To measure level of regret, we asked participants "In your opinion, how serious of a problem was it that you said the messages, at the time you said it" (or tweeted it for Twitter), based on a question from [15]. Participants responded from one (Not at all) to five (Very much so). We consider participants who reported a four or a five to have a high level of seriousness and below a four to have a low level.

Descriptions of Means of Awareness

Self realization	The individual realizes either by thinking about it or by just feeling bad that they should regret the message
Audience says something	The intended audience says something to imply that the person should regret the message
Audience takes an action	The intended audience does something to imply that the person should regret the message (e.g., stops speaking to the individual)
Audience body language	The individual realizes they should regret the message based on the intended audience's body language (e.g., smile, frown)
Third party says something	A person other than the intended audience says something to imply that the person should regret the message
Third party action	A person other than the intended audience does something to imply that the person should regret the message
Third party body language	A person other than the intended audience uses body language to imply that the person should regret the message

Table 2. Codes for means of awareness

For Twitter 18% had high levels of seriousness, while for conversational regrets, 38% had high levels of seriousness. However, the interpretation of the difference is somewhat ambiguous; the seriousness of regrets across contexts may not be directly comparable (e.g., a serious conversational regret may differ from one on Twitter). Levels of seriousness did not vary significantly by type of regret.

Q3: Awareness of regret

Individuals must become aware of a regret to address it. Conversational regret participants tended to become aware of regret more quickly and relied more on audience actions, such as body language cues. Twitter participants more often reported realizing the regret themselves or had audience members tell them they should regret the message.

Means of awareness

We asked each participant to describe, in a free response, how they became "aware you shouldn't have said the message." Two coders created a set of codes for means of awareness based on types of awareness outlined in Meyer's work on regretted messaging [15] using a set of 100 regrets (Table 2). The same two coders then independently coded the regrets based on these codes. A third coder also independently coded the regrets to break ties. In cases where all three coders disagreed, two coders collaboratively reached a consensus. A regret could be coded for multiple, different means of awareness.

Participant-Reported Means of Awareness

	Twitter		Conversation	
Self realization	58%	275	39%	294
Audience said	29%	138	17%	126
Audience action	7%	32	26%	191
Audience body lang	0%	1	19%	143
3rd party said	7%	33	5%	39
3rd party action	1%	5	1%	8
3rd party body lang	0%	1	0%	3
Other	1%	6	0%	3
Total		474		747

Table 3. Means of awareness for Twitter and Conversation

Participants became aware of regret using different means on Twitter and in person (Table 3). This is partially explained by the different contexts for Twitter and conversational regret. Audience body language is usually immediately available in person, but typically absent on Twitter. Thus, 19% of conversational regret participants described using audience body language to become aware of regret. Participants often realized the regret immediately when they saw their audiences' facial expressions. For example, one participant reported calling "his cousin an asshole in-front of our entire family" and realized he should regret it "When everyone glared at me." Conversational regret participants were also more likely to report relying on audience actions to become aware of regret (26% for conversation, versus 7% for Twitter), also likely due to the audience's physical presence. Such actions included storming out of a room, laughter, or sitting silently, which are difficult to convey over Twitter. On Twitter, audience actions that led to awareness sometimes included offline followups, such as laughing about the tweet or the participant being fired, but also included Twitter-specific actions, like being unfollowed, or the intended audience not responding to the regretted message.

Comparatively, Twitter regret participants more frequently became aware of regret on their own (58% versus 39% for conversational regrets). For both, participants would often realize that the regretted message was something that they should not have said or tweeted, either after thinking about it or because they felt bad. As one participant put it: "Something inside just told me it was wrong." However, on Twitter, messages also remain available over time. Several Twitter regret participants reported re-reading the message later and realizing that they should regret it, an option that is rarely available in person. For example, one participant tweeted, "Absolutely pointless," about her relationship and realized she should regret it when she "read over [her] tweets the next morning and thought it was dumb."

Twitter regret participants were also more likely to report that their intended audiences said something to imply that they should regret the message (29% of Twitter, versus 17% of conversational). This partly reflected the wider audiences targeted by Twitter users but also how, on Twitter, people helped participants realize they should regret a message. Often, a

friend or co-worker saw the message and contacted the participant to tell them that they should regret it. For example, one participant tweeted "Having fun on my day off. #callinginsick" and realized he should regret it when "One of [his] friends told [him] it wasn't a good idea."

High or low levels of seriousness of regret did not vary significantly by method of awareness for either Twitter or conversational regrets.

Time until awareness

Conversational regret participants also became aware of regrets more quickly than participants on Twitter. Based on wording used by Meyer [15], we asked participants "how much time passed between" when they tweeted or spoke and when they became aware they shouldn't have tweeted or said the message. We found that the majority of conversational respondents became aware immediately (62%), with many of the remaining participants becoming aware within a few minutes (18%). On Twitter, participants reported taking longer. Only 11% were immediately aware, while 29% realized within a few minutes, 33% at some point the same day, and 16% the next day.

For some types of awareness, participants were more or less likely to become aware immediately (analyses performed with logistic regression). On Twitter, participants were significantly less likely ($p = 0.028$) to become aware of the regret immediately (4%), rather than later, when the audience said something to imply that they should regret the tweet. This is consistent with users tweeting and audience members later informing them that they should regret the content, implying a time delay. For conversational regrets, participants were significantly more likely ($p < 0.001$) to learn immediately (84%) from audience body language about a regret. They often reported realizing as soon as they spoke that they should regret the message due to the audience's physical reactions. As one participant reported "The moment it slipped out, I knew I shouldn't have. The awkward looks and silence that followed confirmed that it was as bad as it sounded." In contrast, conversational regret respondents were significantly less likely ($p < 0.001$) to become aware immediately (13%) when a third party told them something to imply that they should regret the message. The awareness was often delayed because it occurred when a person who they were talking about, or who was impacted by the message, contacted them about it. For example, one participant "told a coworker that I intended to leave my job in an open area" and regretted it "When I went to meet with my boss she told me she had heard rumors."

Q4: Repair strategies

After becoming aware of a regretted message, people will often employ strategies to repair the impact (or potential impact) of the message. We asked participants about the repair strategies they used after tweeting or saying the messages, as well as the impact of these repair strategies. We found that conversational regret participants most often chose to apologize, while Twitter regret participants most often chose to delete the regretted tweet. As occurred in regret awareness, Twitter regret participants also took longer to repair regrets than conversational regret participants.

Participant-Reported Repair Strategies

	Unsuccessful		Successful	
	Twi.	**Conv.**	**Twi.**	**Conv.**
Delete	112	–	136	–
Apology	53	174	72	218
Act like nothing hppnd.	59	81	46	49
Excuse	36	92	34	55
Justify	39	90	30	64
Say something to offset	17	77	22	67
Deny	10	50	10	31
Do nothing	78	109	9	7
Non verbal behavior	–	40	–	31
Other	11	22	5	22
Apology and delete	30	–	38	–
Apology and justify	15	49	16	43
Apology and offset	5	52	12	45
Apology non verbal	–	25	–	19

Table 4. Repair strategies for Twitter and Conversation

Frequency of repair strategy

We asked each participant to select repair strategies they used from a list taken directly from the conversational regrets literature [16]. Both Twitter and conversational regrets participants were provided with the options: "I tried to say something to offset the harm done," "I tried to justify or defend what I said to minimize its offensiveness," "I apologized for saying it," "I just acted like nothing had happened," "I denied or tried to take back what I said," "I offered an excuse for why I said it," "I didn't do anything." Conversational regret participants were also offered the option "I employed a nonverbal behavior to indicate that I regretted it" (from the regrets literature), while Twitter participants were offered "I deleted the tweet."

Overall, we found that a similar proportion of Twitter and conversational regret participants took an action (did not report doing nothing) to repair the regret (82% and 84%, respectively). However, the distribution of repair strategies varied (Table 4). Conversational regret participants most frequently chose to apologize (31% of strategies). Alternatively, Twitter regret participants most often chose to delete the regretted tweet (32%), an option unavailable in conversation. Both conversational and Twitter participants were relatively likely to try to make an excuse (9% of Twitter and 11% of conversational strategies), justify their messages (9% and 12%), and act like nothing had happened (13% and 10%). However, conversational participants were more likely to try to say something to offset the harm (11% of strategies, versus 5% for Twitter).

Success of repair strategies

These different repair strategies also met with varied levels of success (Table 4). Participants rated, on a five point Likert scale, how successful or unsuccessful they felt their choices of repair strategies were. Participants who ranked their strategies as "successful" or "very successful" were categorized as having successfully repaired the regret. Controlling for seri-

ousness of regret at the time of the message, several repair strategies emerged as significantly more likely to be successful or unsuccessful (based on logistic regression).

On both Twitter and in conversation, using an apology significantly increased the probability of successfully repairing harm ($p = 0.043$ and $p < 0.001$ respectively). In person, making an excuse significantly decreased the probability of success ($p = 0.002$), while on Twitter, deleting the tweet significantly increased the probability of successful repair ($p = 0.038$).

Participants who apologized on Twitter varied in their use of online and offline apologies. Online, they apologized using a variety of means, including tweets, instant messages, and text messages. Offline, they apologized face-to-face or by calling impacted individuals. This choice of online or offline strategy seemed to depend on level of personalization and context. Several participants chose to apologize offline because they were confronted about a regretted tweet in an offline environment. For example, one participant apologized when his tennis coach confronted him about an insulting tweet and and told the coach that he "would delete the tweet immediately." Other participants reported apologizing in person to make the apology more personal, writing, "It was personal," so "I called them personally."

Twitter is often a relatively public forum, and, as the regretted tweets often reached wide audiences, apologizing online could also allowed participants to reach larger audiences. Participants reported using online apologies to add additional information to their original tweets or add corrections. One participant, for example, described accidentally posting misinformation about an animal rescue. After realizing her mistake, she tweeted corrected information and an apology. Online apologies were also used to reach large groups of people. One participant described how she "tweeted back so everyone could see my apology and called the person" that she had upset.

Apologies after regretted tweets were also often paired with other online actions. Of the regretted tweets participants apologized for, 54% were also deleted. After posting "something passive-aggressive about someone," one participant described how she tried to repair the situation by telling her "friend that I'd acted immaturely and that I was sorry." She also "deleted the tweet because i was embarrassed by my actions."

For in-person regrets, apologies tended to be offline and verbal, often face-to-face to a single person involved with the regret. As one participant described, he jokingly "insulted a friend only to find out his mother had passed away earlier in the week and hadn't told anyone." Once he found out, the participant "immediately apologized stating that i didn't know and offered my condolences." Such apologies were often paired with justifications (23% of conversational apologies) or explanations that attempted to offset the harm (25%). One participant described criticizing how her husband had done the household chores. She explained that she "apologized, and I think maybe explained that I hadn't meant to sound as rude and critical as it sounded. I also thanked my husband for the work he had done and said that I was glad he was so helpful."

Time to repair

Varied amounts of time passed before participants addressed the regretted messages. Participants responded in free-text to "When did you take these actions?" and one coder coded their responses based on the indication of the first repair (repeated actions were noted once). The categories were: Immediately/a few minutes after the regret (15 or less), the same day, the next day, more than a day but less than a week, more than a week but less than a month, and one month or more. For 30 participants (21 for Twitter and 9 for conversation), the time period was unclear.

Conversational regret participants tended to respond more quickly, as might be expected because they also become aware of the regret more quickly. Of conversational regret participants who tried to repair their regrets, 67% did so within a few minutes. Alternatively, only 28% of Twitter regret participants who tried to repair their regrets did so within minutes; 36% tried to do so the same day, and 19% did so the next day.

LIMITATIONS

There are limitations in our study design. We performed this study using Mechanical Turk. Although this potentially biases our sample, MTurk's population biases have been documented [18]. Samples and results from MTurk workers have also proven comparable to other online sources [6, 7]. We also took several measures to ensure quality responses. However, such quality control measures may also have biased our participant pool, potentially electing for more diligent or intelligent workers. It is unclear how this impact might differ from quality control measures used for other survey methodologies. However, previous conversational regrets work drew from an undergraduate population [15, 16]; using MTurk allowed us to expand to a large, cost-effective sample relative to both such offline pools and alternative online sources.

Our survey design had additional, inherent limitations. We used the basic design from the conversational regrets literature [15, 16] in which each participant recalled a single, regretted message. Thus, we didn't have a true analysis of the frequency of different types of conversational or Twitter-based regrets. Based on the conversational regrets design, we asked participants for the regret that first came to mind, rather than the most recent or strongest regrets. However, certain regrets may come to mind more easily or may be more or less embarrassing to detail in a survey. Thus, we may have an overrepresentation of memorable regrets and an underrepresentation of deeply shameful regrets.

The survey format was also a limitation. We asked participants for self-reported, recalled data. Participants may attribute more meaning to events occurring in the past when reporting on them in a survey. There was also potential for reverse causality issues. We tried to limit causality questions, but participants may have attributed factors like states of being to the regret, when they were actually caused by the regret. We could offer more conclusive results if we tracked

participant behavior over time and noted actions, like repair strategies, as they occurred. Alternatively a diary-study approach could be used to supplement this work.

DISCUSSION

We found that Twitter and conversational regret participants differed in the types of messaging regrets they reported, how they became aware of the regrets, and how they tried to repair the harm caused by the regrets. Time delays on Twitter, as well as lack of face-to-face communication with audiences, also caused awareness and repair on Twitter to occur more slowly than it did for conversational regrets. Based on these findings, we offer several early potential design directions for helping users prevent and repair Twitter regret.

Detecting and preventing regret on Twitter

Although our participants took measures to repair harm caused by the regretted messages, they often would have liked to have not tweeted the messages. One way to potentially prevent regret on Twitter would be to develop tools to detect potentially regrettable messages and provide users with suggestions for when they might want to reconsider tweeting. Behavioral economics offers a potential direction to help prevent users from sending such tweets by using behavioral "nudges" to help people identify tweets they might not want to post [1, 4]. Such nudges are cues that suggest that users should alter a behavior without forcing them to do so.

We found that several negative emotions, including anger, stress, and frustration, tended to lead to regret on Twitter. A recent study of deleted tweets also found a slightly higher frequency of negative-sentiment keywords in tweets that were deleted [3]. As we found, deletion is a common strategy for coping with regretted tweets. Prior to a tweet being sent, such negative states could potentially be detected using tools like sentiment analysis or word frequency. Word analyses could potentially also be combined with environmental cues, such as location, especially when users tweeted from mobile devices; 45% of regrets reported by Twitter regret participants were made from mobile devices. Once a negative mood was detected, it might be possible to provide feedback to the user about the negative emotion, or, in a manner similar to Google Mail Goggles [17], lock them out until they could think more clearly.

We also found that certain types of regret related to broadcasting thoughts to wide audiences were more common on Twitter. Twitter regret participants tended to report regretting revealing too much, revealing expressive/cathartic thoughts, and sharing with unintended audiences. Such types of regret might be preventable through better audience awareness or management on Twitter. Participants often regretted tweets that revealed too much or that were expressive/cathartic because they were seen by people they didn't want to see them, or because people saw the tweets and were hurt. For these regrets, it might be possible to more clearly indicate who might view a tweet, for example by showing images of a user's followers. Interestingly, several tweets were sent by participants who had protected accounts at the time of the regretted message (25% overall, and 21% for unintended audience). Participants tended not to accidentally tweet to the general public. Rather, their tweets were viewed by people they didn't initially anticipate would view the posts. This is in line with Acquisti and Gross' concept of "imagined communities" [2] and Marwick and boyds concept of tweeting to an "imagined" audience [13]. One possible way to visualize the actual audience would be to show images of people who could view the tweet, potentially prioritizing them by interaction level. For instance, Lieberman and Miller's Facemail prototype uses this approach for email [11].

Promoting regret awareness

To address a regretted message, users must first realize that they should regret the tweet. We saw several methods for becoming aware of regretted messages that were unique to in-person conversation and could potentially be adapted for Twitter, as well as several techniques that were unique to Twitter and could be further emphasized.

In person, participants often quickly become aware of regrettable messages, typically through physical cues from audiences, like body language. It was easy for one participant to tell that he should regret the message after his girlfriend "instantly became upset and started to cry." Other participants saw audiences storm out of the room or laugh. Twitter users, physically separated from their audiences, usually lack these sources of instant audience feedback.

One possibility for improving Twitters users' awareness of regret would be to improve their abilities to gauge potential audience reaction absent physical feedback. Work has been performed to visualize sentiment conveyed in electronic communications. For example, Liu et al. prototyped an "EmpathyBuddy" for email that presents a line-drawn face that reacts to the emotion in the text [12]. Similar visualizations showing the sentiment conveyed by tweets might help Twitter users more quickly become aware of potentially regrettable tweets before tweeting them. A visualization that persisted after a user tweeted might also allow awareness to occur more quickly after a tweet.

We also found that, on Twitter, participants often reported being informed by their communities (e.g., friends, family, and co-workers) that they should regret messages, often over electronic means like text messages or Twitter itself. Lampinen et al. discussed how users of social networks collaboratively control disclosure [9]. Similar to Lampinen et al.'s participants, who used collaborative strategies to protect each others' privacy, other people helped our participants become aware of regretted content both when it affected the audience (or a third party) and when it only affected the participant. Developing easy mechanisms for people to tell someone about potentially regrettable tweets could mitigate potential regret.

Throughout our results, we saw that Twitter had a large time delay compared to conversation, both in terms of time to awareness and time to repair. This was somewhat due to the lack of immediate audience feedback; in cases where Twitter regret was informed by others, this response often came hours or days later. On Twitter users both cannot see typically see immediate feedback, and audiences sometimes can-

not immediately access messages, delaying regret awareness and potential repair. However, unique to Twitter, even when there was no negative reaction, participants regretted tweets because of the record provided by Twitter. Participants re-read their tweets and realized the message was regrettable. Creating tools that better help users review past tweets may also help them become aware of, and purge, possibly regretful content.

CONCLUSION

We examined Twitter users' regrets for in-person conversations and on Twitter. We found that, on Twitter, participants tended to report more regrettable messages targeted at broad audiences, including messages intended to be expressive or cathartic, that revealed too much, or that reached unintended audiences. In general, we also saw that Twitter regret participants became aware of regret more slowly than conversational regret participants, more often relying on others to tell them about the regret or eventually realizing themselves that the message should be regretted in the absence of physical audience cues. Once aware of the regret, Twitter users tended to delete the regretted tweet and/or apologize. Based on the findings, we offer several early design suggestions, including behavioral nudges for helping Twitter users realize potentially regrettable posts either before or after tweeting, and for better audience management.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grants No. 0946825, DGE-0903659, and CNS-1012763 (Nudging Users Toward Privacy), by Google under a Focused Research Award on Privacy Nudges, by IWT, by a DoD NDSEG Fellowship and by the ARCS Foundation.

REFERENCES

1. Acquisti, A. Nudging privacy: The behavioral economics of personal information. *IEEE Security & Privacy 7*, 6 (Nov./Dec. 2009), 82–85.
2. Acquisti, A., and Gross, R. Imagined communities: Awareness, information sharing, and privacy on the facebook. In *Privacy enhancing technologies*, Springer (2006), 36–58.
3. Almuhimedi, H., Wilson, S., Liu, B., Sadeh, N., and Acquisti, A. I wish I hadn't tweeted that! Large-scale quantitative analysis of deleted tweets. In *Proc. CSCW 2013*, ACM (2013).
4. Balebako, R., Leon, P., Almuhimedi, H., Kelley, P. G., Mugan, J., Acquisti, A., Cranor, L. F., and Sadeh, N. Nudging users towards privacy on mobile devices. CHIPINC 2011 (2011).
5. Bosker, B. The Twitter typo that exposed Anthony Weiner, 2011. `http://www.huffingtonpost.com/2011/06/07/anthony-weiner-twitter-dm_n_872590.html`.
6. Buhrmester, M., Kwang, T., and Gosling, S. D. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science 6*, 1 (2011), 3–5.
7. Jakobsson, M. Experimenting on Mechanical Turk: 5 how tos. `http://blogs.parc.com/blog/2009/07/experimenting-on-mechanical-turk-5-how-tos/`, July 2009.
8. Knapp, M. L., Stafford, L., and Daly, J. A. Regrettable messages: Things people wish they hadn't said. *Journal of Communication 36*, 4 (1986), 40–58.
9. Lampinen, A., Lehtinen, V., Lehmuskallio, A., and Tamminen, S. We're in it together: Interpersonal management of disclosure in social network services. In *Proc. CHI 2011*, ACM (2011), 3217–3226.
10. Lampinen, A., Tamminen, S., and Oulasvirta, A. All my people right here, right now: management of group co-presence on a social networking site. In *Proc. GROUP 2009*, ACM (2009), 281–290.
11. Lieberman, E., and Miller, R. C. Facemail: showing faces of recipients to prevent misdirected email. In *Proc. SOUPS 2007*, ACM (2007), 122–131.
12. Liu, H., Lieberman, H., and Selker, T. Automatic affective feedback in an email browser. Tech. rep., MIT Media Laboratory Software Agents Group, 2002.
13. Marwick, A. E., and boyd, d. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society 13*, 1 (Feb. 2011), 114–133.
14. McLaughlin, M. L., Cody, M. J., and O'Hair, H. D. The management of failure events: some contextual determinants of accounting behavior. *Human Communication Research 9*, 3 (1983), 208–224.
15. Meyer, J. R. Regretted messages: Cognitive antecedents and post hoc reflection. *Journal of Language and Social Psychology 30*, 4 (2011), 376–395.
16. Meyer, J. R., and Rothenberg, K. Repairing regretted messages: Effects of emotional state, relationship type, and seriousness of offense. *Communication Research Reports 21*, 4 (2004), 348–356.
17. Perlow, J. New in labs: Stop sending mail you later regret, 2008. Official Gmail Blog. `http://gmailblog.blogspot.com/2008/10/new-in-labs-stop-sending-mail-you-later.html`.
18. Ross, J., Irani, L., Silberman, M. S., Zaldivar, A., and Tomlinson, B. Who are the crowdworkers?: Shifting demographics in Mechanical Turk. In *Ext. Abstracts CHI 2010*, ACM (2010), 2863–2872.
19. Wang, Y., Norcie, G., Komanduri, S., Acquisti, A., Leon, P. G., and Cranor, L. F. "I regretted the minute I pressed share": A qualitative study of regrets on Facebook. In *Proc. SOUPS 2011*, ACM (2011).
20. Wisniewski, P., Lipford, H., and Wilson, D. Fighting for my space: Coping mechanisms for SNS boundary regulation. In *Proc. CHI 2012*, ACM (2012), 609–618.